2007

An Investigation Of The Impact On Student Achievement Of The Mobile Math Initiative In Three High Poverty Schools In A Single School District

Jeanne D. Payne
University of Alabama at Birmingham

AN INVESTIGATION OF THE IMPACT ON STUDENT ACHIEVEMENT OF THE MOBILE MATH INITIATIVE IN THREE HIGH POVERTY SCHOOLS IN A SINGLE SCHOOL DISTRICT

by

JEANNE D. PAYNE

JERRY PATTERSON, COMMITTEE CHAIR
GYPSY ABBOTT
LOUCRECIA COLLINS
MARGARET RICE
BOYD ROGAN

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2007

AN INVESTIGATION OF THE IMPACT ON STUDENT ACHIEVEMENT OF THE MOBILE MATH INITIATIVE IN THREE HIGH POVERTY SCHOOLS IN A SINGLE SCHOOL DISTRICT

JEANNE D. PAYNE

EDUCATION LEADERSHIP

ABSTRACT

Teachers must increase effectiveness of teaching to prepare quantitatively literate students. The study's purpose was to determine the effectiveness of the Mobile Math Initiative (MMI) in three high poverty schools. Changes in achievement of 89 students over 3 years were examined using the Stanford 10 (SAT 10) Procedures and Problem Solving subtests and the Alabama Reading and Mathematics Test (ARMT) Math subtest. Independent variables were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended. Data were analyzed with the SPSS (Version 15.0) software using factorial ANOVA with repeated measures.

For the SAT 10 Procedures subtest, the null hypothesis of no significant interaction for school was rejected ($F = 7.78$, $df = 4$, $p < .001$). There was also significant interaction between scores and testing occasions ($F = 25.57$, $df = 2$, $p < .001$). For SAT 10 Problem Solving subtest scores, the null hypothesis of no significant interaction for school was rejected ($F = 3.35$, $df = 4$, $p = .011$). There was also significant interaction between scores and testing occasions ($F = 11.18$, $df = 2$, $p < .001$). For the ARMT Math subtest, the null hypothesis of no significant interaction for school was rejected ($F = 3.43$, $df = 2$, $p = .037$). For all subtests, independent *t* tests were only significant for special education. Paired *t* tests comparing mean NCE scores for the SAT 10 Procedures subtest for each category of independent variable at baseline and final testing were significant for

all levels of all independent variables except White and special education students. Paired *t* tests comparing mean NCE scores for the SAT 10 Problem Solving subtest for each category of independent variable at baseline and final testing were significant for all levels of all independent variables except White, paid lunch, and special education students. Paired *t* tests comparing mean scale scores for the ARMT Math subtest for each category of independent variable between testing occasions were significant only for special education students. The researcher concluded that the MMI was effective in high poverty schools.

ACKNOWLEDGMENTS

As I conclude this product of data, methods, and research, I find myself reflecting on individuals who have supported and nurtured my mind and heart. Many people have helped expand and educate my mind.

Members of my committee –Jerry Patterson, Gypsy Abbott, Loucrecia Collins, Margaret Rice, and Boyd Rogan – endeared me with their personal touches of support, guidance, and encouragement throughout my journey through the Educational Leadership program. Dr. Patterson accepted nothing less than high standards (which was not easy, but it sure felt good when the mission was accomplished!). I appreciate his willingness to chair my committee and, as well as his excitement over the findings of my study. Dr. Abbott spent many hours working with me—even on weekends. I treasure her friendship. Dr. Collins made a special trip to Decatur, Alabama, to sit with my family for an award ceremony. I value the lessons I learned in her class. Dr. Rice, who is always positive and helpful, reassured me when I caught a terrible virus and had to miss her class during a crucial week. Dr. Rogan, who has kind words for everyone, gave a beautiful eulogy that touched my heart at Dr. Harold Bishop's funeral. You intertwined life lessons with scholarly precepts, encouraging me to stretch, academically.

All of my teachers, who I respect and revere, afforded me knowledge and opportunities to help satisfy my insatiable appetite for learning. You made me more confident, courageous, and compassionate, as I found my place in the world.

The University of Alabama at Birmingham has played a large part in my academic and personal development. I have now finished five college degrees, three of which are from UAB. My studies here were interesting and challenging. More than 20 years ago, a UAB teacher casually mentioned Viktor Frankl's *Man's Search for Meaning*; his story has influenced my life. Also many years ago, I happened to take three one-hour classes on women's studies; I realized endless opportunities were available for me. Many education classes closely correlated with the work I was doing in schools, enhancing my skills through high-quality instruction. I appreciate the efforts of Dr. Foster Watkins, who encouraged me and many other students to continue our studies at UAB.

A diverse group of very bright friends sustained me at UAB. Donna Burnett, of the Greater Birmingham area, not only edited my paper with her keen mind and vast knowledge of APA style, but offered me moral support and true friendship (she has a way of making long hours of work enjoyable). I predict big things for Donna as she completes her doctoral studies. Moayad Wahsheh, from, Jordan, has a delightful family and is a prefect gentleman. Grace Jepkemloi, with her beautiful smile, will probably be president of Kenya someday. Ramsey Huffman, my co-worker and college buddy, often talked continuously with me on the countless commutes back to Decatur from UAB, to keep me from falling asleep while driving; our good friendship, determination, and hard work have gotten us this far, and I believe he will complete his studies soon. For my good friends and others who have enriched my life, I am thankful because achievement does not happen through solo efforts. We are all interdependent.

Dr. Susan Pruet, the mother of the Mobile Math Initiative, has boosted student and teacher understanding of math through her brilliant efforts and perseverance. I called her many times, asking her to please teach us what she was doing in mathematics instruction. My persistence paid off, and Susan began working with us. She also introduced me to Dr. Honi Bamberger, a highly gifted teacher of teachers. The success of the Mobile Math Initiative is directly related to their efforts. I hope this dissertation does justice to their work. Also important to the success of the MMI is the hard work of mathematics coaches Linda D'Antonio, Leah Edgecomb, Terri Patterson, and Connie Teague, and the efforts of MMI principals and teachers.

Fine folks, particularly my colleagues in Decatur City Schools, have enhanced my professional life in education. Decatur City Schools is a system where people are trained in Stephen Covey's *7 Habits of Highly Effective People* to create a climate of competence, collaboration, personal trustworthiness, and emotional bank account deposits. It has been a pleasure to work in a school system that devotes resources for high quality, sustained, and data-driven professional development as the best way to improve student achievement.

Dr. Sam Houston, a gifted superintendent and curriculum expert with an incredible ability to solve problems, is a people person who values the worth of each employee and student in his school system. Dr. Larry Walters, a smart leader who quietly got things done, and Dr. Buddy Davis, an energetic change agent, were my former superintendents. All three men believed in my abilities, gave me wonderful and challenging job opportunities, and became important role models. Phil Hastings, for whom I did my student teaching (at the time, I thought he was working me to death!), gave me a solid foundation. Dr. Tim Lull, who, when he was at Decatur High School, gave me my first administrative job as an assistant principal; Dr. Lull trained me in every possible area of school administration with his Army-laced orders: "Payne, do this," or "Payne, do that." He believed I could do any task – no matter how hard – and I grew to believe that, too.

After Tim's "boot camp," being a middle school principal was a piece of cake! Other talented people have been great encouragers, including Dr. Ed Nichols, Melanie Maples, Dr. Yvette Evans and Sandra Calvin, always upbeat and full of ideas, and supervisors, principals, assistant principals, and school employees. I also appreciate the support of my talented friends, Jan Mendenhall and Dr. Carol Chenault. Billie Patton, my smart, skilled, positive, and funny assistant, is my greatest encourager and a dear friend. I also appreciate the encouragement of Jennie Adams, Sole Gonzalez and Joyce Ratley.

Throughout my life, my family has nurtured my heart. My husband Gary has been by my side since the beginning of my career, when my first teaching job was a far drive out into the country. Your love, encouragement, and sense of humor – and the long walks and boat rides at sunset – have been invaluable to me.

My sons, Adam and Blake, make life sparkle for me. You add never-ending joy and an indescribable dimension to my life. I was proud of Adam the day he graduated from college in marketing, and I am equally proud of Blake as he pursues aerospace studies. Adam has a wonderful ability to transact business smoothly and is a whiz with technology. Blake is a gifted conversationalist. Blake can fix anything and make a yard look like it belongs on the cover of a magazine. Exceptions to this joy and pride include my finding Blake, then 3 years old, in the middle of the road directing traffic around a dead squirrel; finding the dog chewing a hole in the living room floor when I was assured by the boys that they had not let the dog in; being assured, along with Dad, that a college student and a high school senior did not need a babysitter while their parents were out of the country; and finding out about Adam getting a speeding ticket in the "Nava Joe" and going to driving school without telling us—all because he thought we would be angry (we were just impressed that the car would travel over 50 miles per hour!). In spite of having a mother who works in schools, loves to go to school, and talks about school, you have grown into fine, productive young men who are a great source of pride for our family. I appreciated your encouragement as I attended school and did homework.

My mother, Doris Duke, read poetry to me daily when I was a preschooler and later directed me to the awe-inspiring World Book Encyclopedia. She encouraged me, believed in my abilities, and, through all events of life, gave me the security of unconditional love. She also passed along her positive-attitude gene, which has made my life's journey much more than just "tolerable." My late father, Twigg Duke, instilled in me a relentless work ethic. He had a silver bowl given to him as a gift; as a child I ran my fingers over the letters on the inscription: "A Great Leader." It made me think about being a leader.

My brother and sisters – David Duke, a gifted businessman; Susan Chenault, a gifted social worker; and Emily Payton; a gifted teacher – continue to give me love and support. Their delightful spouses – Karen Duke, Marty Chenault, and Wally Payton – and all the nieces, nephews and cousins energize me. I look forward to sharing many more birthday parties with you. You all have been wonderful examples for me as you have pursued excellence in your own lives. I wonder if, as a child, the many hours of playing school with my sister Susan influenced my decision to become a teacher.

Two very special, influential women passed away while I was working on my doctorate. Charlotte Lake, my mother-in-law, was a marvelous cook who loved to entertain. Her house was the only one I knew of where someone was prepared to feed a family of five, who unexpectedly dropped by at mealtime. Oh, how I miss that potato salad! She always treated me like a daughter and showed me love, kindness, and encouragement. My aunt, Betty Faye Barton, adored her yard with its crystal pink camellias, blue pansies with faces, heavenly sweet olive, and moss-laden Selma trees. Betty Faye, with her impeccable taste and quiet activism, had a way of making each family member feel special. The *Jeannie* book she wrote filled many children with joy, including me, and has a place of honor in my home.

My family has long valued education, recognizing that with knowledge comes the responsibility to share with others. I think of my grandmother, Mama Jean Barton, who, as an immigrant child, hid in the back of Carnegie Hall to listen to beautiful music; my grandpa, Solomon Barton, who read everything known to man; my aunt, Evelyn Weisfeld, of New York, who is the renaissance woman and teacher; my uncle, Albert "Bubber" Barton, who was named after Albert Einstein, of course; and my aunt, Dot McCain, who is an encourager to everyone.

Someday I hope to hold my grandchildren and great-grandchildren in my lap and tell them about my passion for life. If they balk at hearing me read pages of this dissertation, perhaps they will listen to what I have learned: work hard, do the right thing, learn all you can, say thank you, be passionate about what you do, and love with all of your heart.

Muhammad Yunis, when he accepted the 2006 Nobel Peace Prize, said, "A human being is born into this world fully equipped not only to take care of him or herself, but also to contribute to the enlarging well-being of the world as a whole. Some get the chance to explore their potential to some degree, but many others never get any opportunity, during their lifetime, to unwrap the wonderful gift they were born with. They die unexplored and the world remains deprived of their creativity, and their contribution." My hope is that by acquiring a doctorate from UAB, the knowledge I gain will enable me to help others explore their potential.

Some people search a lifetime to find meaningful work. I am thankful that I awaken every morning, excited about educating children. Hans Selye said, “A long, healthy, and happy life is the result of making contributions, of having meaningful projects that are personally exciting and contribute to and bless the lives of others.” I appreciate the opportunity to fulfill my educational dreams and to satisfy my hunger for learning. These experiences have made me feel incredibly alive.

God, the author and perfecter of my faith, has cared for and nourished my soul and given me strength for this day. I am eternally grateful.

TABLE OF CONTENTS

Page

ABSTRACT iii

ACKNOWLEDGMENTS v

LIST OF TABLES xv

LIST OF FIGURES xviii

CHAPTER

1 INTRODUCTION 1

Statement of the Problem 6
Purpose of the Study 7
Null Hypotheses 7
Significance of the Study 8
Definition of Terms 9
Assumptions 12
Limitations 12
Organization of the Study 13
Summary 13

2 REVIEW OF LITERATURE 14

Mathematics Reform Efforts 16
National Mathematics Studies 16
National Assessment of Education Progress 17
International Mathematics Studies 19
Program for International Student Assessment 19
Third International Mathematics and Science Study and Trends in International Mathematics and Science Study 20
U.S. Efforts to Improve Mathematics Achievement 21
National Council of Teachers of Mathematics Standards 21
The National Science Foundation 23

TABLE OF CONTENTS (Continued)

Page

CHAPTER

Promising and Exemplary Mathematics Programs 24
Promising Mathematics Programs 25
An Overview of the MMI 27

3 METHODOLOGY 33

Intervention 35
Levels of Personnel Involved with MMI Implementation 39
Teacher Leadership Program 40
Implementation Committee 40
Sample 41
Data Collection Procedures 44
Data Analysis Plan 45
Instruments 46
ARMT Validity and Reliability 46
Validity of the ARMT 46
Reliability of the ARMT 47
SAT 10 Validity and Reliability 48
Summary 49

4 FINDINGS 51

Introduction 51
Data Verification 51
Procedures for Repeated Measures 52
Assumptions of Repeated Measures 53
Demographic Characteristics of the Samples 53
Findings of Statistical Analysis for Null Hypothesis 1 56
Statistically Significant Findings for Null Hypothesis 1 57
Conclusions for Null Hypothesis 1 66
Findings of Statistical Analysis for Null Hypothesis 2 67
Statistically Significant Findings for Null Hypothesis 2 68
Conclusions for Null Hypothesis 2 78
Findings of Statistical Analysis for Null Hypothesis 3 79
Statistically Significant Findings for Null Hypothesis 3 80
Conclusions for Null Hypothesis 3 89
Summary 89

TABLE OF CONTENTS (Continued)

Page

CHAPTER

5 SUMMARY, DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS....91

Summary 91
MMI Professional Development....95
Findings of the Null Hypotheses....100
Discussion....110
Implications....113
Conclusions....116
Recommendations for Improving the Research....117
Recommendations for Future Research....118

LIST OF REFERENCES....121

APPENDIX

A THE MOBILE MATH INITIATIVE 126

B INSTITUTIONAL REVIEW BOARD FOR HUMAN USE (IRB) APPROVAL 129

LIST OF TABLES

Table *Page*

1 MMI Teacher Training Schedule..35

2 MMI Big Ideas..36

3 MMI Professional Development Training Progression 39

4 Combined Demographics for Elementary Schools A, B, and C by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education needed 42

5 Elementary School A Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 42

6 Elementary School B Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 43

7 Elementary School C Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 43

8 Data Collection Schedule for Standardized Testing 44

9 Intercorrelations of Domains, Subdomains and Total Scores for the ARMT Math Subtest for Grade 3 47

10 Interrater Agreement Coefficients for the ARMT Math Subtest for Grade 3.............48

11 Reliability Coefficients for the SAT 10 Mathematics Subtests for Grade 3...............49

12 Combined Demographics for Elementary Schools A, B, and C by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 54

13 Elementary School A Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 55

14 Elementary School B Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 55

LIST OF TABLES (Continued)

Table *Page*

15 Elementary School C Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 56

16 Descriptive Statistics for NCE Scores for the SAT 10 Procedures Subtest for All Testing Occasions by School 57

17 Multivariate Tests for NCE Scores for the SAT 10 Procedures Subtest 58

18 Descriptive Statistics for NCE Scores for the SAT 10 Procedures Subtest for Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed for All Testing Occasions 58

19 Independent *t* Tests for NCE Scores for the SAT 10 Procedures Subtest for Gender, Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed 60

20 One-Way ANOVA for Ethnicity for NCE Scores for the SAT 10 Procedures Subtest 60

21 Paired *t* Tests for NCE Scores for the SAT 10 Procedures Subtest Between 2004 and 2006 for Levels of Ethnicity, Gender, Level of Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 61

22 Paired *t* Tests for NCE Scores for the SAT 10 Procedures Subtest Between 2004 and 2005 and Between 2005 and 2006 by School 61

23 Descriptive Statistics for NCE Scores for the SAT 10 Problem Solving Subtest for All Testing Occasions by School 67

24 Multivariate Tests for NCE Scores for the SAT 10 Problem Solving Subtest 68

25 Descriptive Statistics for NCE Scores for the SAT 10 Problem Solving Subtest for Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed for All Testing Occasions 69

26 Independent *t* Tests for NCE Scores for the SAT 10 Problem Solving Subtest for Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 72

27 One-Way ANOVA for Ethnicity for NCE Scores for the SAT 10 Problem Solving Subtest 72

LIST OF TABLES (Continued)

Table *Page*

28 Paired *t* Tests for NCE Scores for the SAT 10 Problem Solving Subtest Between 2004 and 2006 for Levels of Ethnicity, Gender, Level of Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 73

29 Paired *t* Tests for NCE Scores for the SAT 10 Problem Solving Subtest Between 2004 and 2005 and Between 2005 and 2006 by School 73

30 Descriptive Statistics for Scale Scores for the ARMT Math Subtest 80

31 Multivariate Tests for Scale Scores for the ARMT Math Subtest 80

32 Descriptive Statistics for Scale Scores for the ARMT Math Subtest for Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed for Both Testing Occasions 81

33 Independent *t* Tests for Scale Scores for the ARMT Math Subtest for Gender, Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed 83

34 One-way ANOVA for Ethnicity for Scale Scores for the ARMT Math Subtest..83

35 Paired *t* Tests for Scale Scores for the ARMT Math Subtest Between 2005 and 2006 for Levels of School, Ethnicity, Gender, Level of Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed 84

36 Combined Demographics for Elementary Schools A, B, and C by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed 99

LIST OF FIGURES

Figure *Page*

1 NCE scores for the SAT 10 Procedures subtest across testing occasions 59

2 NCE scores for the SAT 10 Procedures subtest reported by school............................62

3 NCE scores for the SAT 10 Problem Solving subtest across testing occasions 74

4 NCE scores for the SAT 10 Problem Solving subtest reported by school...................75

5 Scale scores for the ARMT Math subtest between testing occasions..........................82

6 Scale score for the ARMT Math subtest reported by school 85

CHAPTER 1

INTRODUCTION

According to the No Child Left Behind (NCLB) Act (2001), all students are required to be proficient in mathematics by 2014. Achieving this goal requires exploring the methods that teachers are currently using to teach mathematics and assessing whether these methods are increasing achievement for all students. Because many researchers report that teacher quality is the most important factor in student achievement, high-quality professional development for currently practicing teachers and preservice teachers becomes a key element in giving teachers the skills needed to bring all students to the proficiency level in mathematics (Darling-Hammond & Sykes, 1999).

The American Institutes for Research (2006) reported that, in 2003, "20% of U.S. college students completing 4-year degrees—and 30% of students earning 2-year degrees—have only basic quantitative-literacy skills" (p. 1). *Quantitative literacy* has been defined as the possession of the knowledge and skills required to perform real-world tasks such as completing an order form, balancing a checkbook, or figuring out how much interest would be due on a loan (White & McCloskey, 2005). Steen (1999) purports that, when individuals think about what they are doing and understand why they are doing it, they are far more successful than those are who follow the rules without understanding. When students learn to think logically and with understanding—requirements of quantitative literacy—they will be better prepared for a more productive life.

In addition, 58% of students in the 2006 graduating class who took the ACT assessment received a score that indicated that they were not ready for college-level algebra. These data suggest that too many students are being left behind in terms of learning mathematical concepts and skills. The most plausible explanation is that there are gaps in the teaching and learning of mathematics in kindergarten through 12th grade and that these gaps are leaving many students unprepared for the next level of study (Fosnot & Dolk, 2001). In the United States., a majority of students apparently do not have a deep understanding of mathematics, have not developed their capacity to think logically, and are unable to apply their knowledge of mathematics to real-life situations. If teachers have not been taught the skills necessary for quantitative literacy, it is not likely that they can teach these skills without significant professional development.

Smith (2001) has concluded that educators should not be content with minimal reform but must insist on a true overhaul in their thinking about mathematics. Thompson and Zeuki (as cited in Darling-Hammond & Sykes, 1999) have suggested that the kind of learning that will be required of teachers should be described as transformative and will involve major changes in deeply held beliefs, knowledge, and habits of practice. Learning cannot simply be the addition of new skills to the existing repertoire. In 2003, the U.S. Department of Education sponsored the Secretary's Summit on Mathematics. At the summit, speaker Deborah Loewenberg Ball took the idea of professional development to a deeper level as it applies to both preservice and in-service teachers. Ball (2003) stated that

> the quantity of teachers' mathematics coursework will only improve the quality of mathematics teaching if teachers learn mathematics in ways that make a difference for the skill with which they are able to do their work. The goal is to improve students' learning. Teachers' opportunities to learn must equip them with the mathematical knowledge to teach mathematics effectively. (para. 3)

Ruth Parker (2005), CEO of the Mathematics Education Collaborative (MEC), stated that the goal of mathematics education is to enable students to be mathematically confident and competent. She went on to say that mathematics confidence comes from "knowing that you understand mathematics and that you have the tenacity to make sense of information and situations that you encounter even when the problems are complex or messy" (Parker, para. 1). These learning opportunities for teachers are central tenets of the Mobile Math Initiative (MMI), which offers teachers in Grades K-5 in-service training designed to transform their learning about mathematics. It is anticipated that the specific professional development provided through the MMI will lead to effective teaching practices that prepare students to be quantitatively literate citizens of the world.

Findings from Leinwand (2000) suggest that a student's mathematical problem-solving ability is a reflection of the mathematical learning culture created by his or her teacher. When examining the methods that have been used to teach mathematics, researchers have shown an apparent disconnect between procedural and conceptual mathematics during the past 20 years (Leinwand). Even after the establishment of state and national standards, the quality of mathematical knowledge possessed by students has changed only minimally since the National Council of Teachers of Mathematics (NCTM) standards were first introduced in 1989. According to the 2003 National Assessment of Education Progress (U.S. Department of Education, 2003), 31% of eighth-grade students still scored below the basic level in mathematics. Many students follow the rules and carry out procedures that they do not understand; this approach makes it difficult for them to modify their skills to fit new situations.

When teaching methods in general are considered, it becomes apparent that there are some firmly embedded ideas in American culture about knowledge and learning. One idea is that knowledge is fixed and that teachers transmit knowledge to pupils who store it and remember it. As many as 300 years ago,

> most teaching proceeded as though learning was a passive process of assimilation. Students were expected to follow their teachers' directions rigorously. The study was to imitate: to copy a passage, to repeat a teacher's words, or to memorize some sentence, dates, or numbers. Students may have posed questions in informal discourse, and perhaps even embroidered the answers. But school learning seems to have been a matter of imitative assimilation. (Cohen, 1989, pp. 42-43)

In the American educational system, there has been a push toward basic skills and the computation and memorization of facts. This push is a direct reflection of the accountability required in the NCLB Act (2001). There appears to be a lack of incentives to work on complex content or to delve into the looseness of an exploratory curriculum. "The pull toward neat, routinized instruction is very strong" (Ball, Lubienski & Mewborn, 2001, p. 436). Teachers often work in isolation, with a lack of sufficient time to collaborate with each other and a lack of time for support of teacher learning (Ball et al., 2001).

Reviews of the way mathematics is taught in the international arena have revealed that multiple methods are being used, with some methods resulting in higher levels of students' mathematics knowledge. Furthermore, some of these methods are leading to higher student achievement in mathematics. In the Third International Mathematics and Science Study (TIMSS), a videotaped classroom study of 1994-1995 (Stigler & Hiebert, 1997), eighth-grade classrooms from the United States and Japan were compared. Strategies used by U.S. teachers and teachers in Japan were dramatically different. According to the TIMSS report (Stigler & Hiebert, 1997), findings indicated that the purpose of Japanese lessons was to enhance students' understanding of mathematical concepts. This approach is in contrast to the instructional methods typically used in U.S. classrooms. In comparison with their Japanese counterparts, American teachers were reported as spending less time presenting new material and most of their time reviewing topics that had already been taught. American teachers asked mostly yes/no questions about the work. American students sat in rows and practiced mathematical problems that they had already been shown how to do. Thus, students were not required to apply mathematics skills in novel situations.

The following quotation provides a description of the teaching found in the TIMSS (Stigler & Hiebert, 1997) videotape:

> The typical eighth-grade mathematics lesson in the U.S. is organized around two phases: an acquisition phase and an application phase. In the acquisition phase, the teacher demonstrates or leads a discussion on how to solve a sample problem. The aim is to clarify the steps in the procedure so that students will be able to execute the same procedure on their own. Students in the application phase practice using the procedure by solving problems similar to the sample problem. (p. 18)

Analysis of the videotape provided some powerful insights. In the videotape, 78% of the topics or concepts taught were not fully developed by the American teachers, who presented procedures rather than concepts. In 96% of the case studies, students were engaged in seat work; they practiced procedures. Only 1% of the students were doing more in-depth mathematics thinking such as analyzing new problems or developing alternative solutions. Notable in their absence were several strategies for teaching mathematics that provided opportunities to (a) solve challenging problems, (b) reason mathematically, and (c) communicate with intent to justify (Stigler & Hiebert, 1999).

According to results of research by Liping Ma (1999), Chinese students regularly outperform U.S. students on international assessments of mathematics competency, yet

Chinese teachers have been described as having less mathematics training than U.S. teachers do. Chinese teachers have 11 or 12 years of formal education, whereas U.S. teachers have 16 to 18 years of school before they are certified to teach. Ma contends that this paradox occurs because Chinese teachers have a much deeper understanding of elementary mathematical concepts than their U.S. counterparts do. This disparity seems to suggest differences in preservice education, as well.

What is the current status of mathematics instruction? A good source of information is the 2003 version of the NAEP (U.S. Department of Education, 2003). In this study of high school seniors who took the test, 63% could not figure out a simple multiplication problem in which calculation of the amount of postage was needed for a package of a given weight. Even with the use of a calculator, 88% of seniors taking the 2003 NAEP (U.S. Department of Education, 2003) could not calculate the interest that they would earn on a savings account. If students cannot do these tasks, one can only assume that they did not receive adequate instruction from the teachers who facilitated this instructional program.

Statement of the Problem

The U.S. educational system has repeatedly implemented reform efforts in the teaching of mathematics, with very little change in classroom performance (Leinwand, 2000). According to analysis and conclusions from the TIMSS video study (Stigler & Hiebert, 1997), U.S. teachers often use traditional teaching methods to show students what to do, and the teachers do not develop specific mathematics concepts to foster a deeper level of understanding. Problems assigned for practice are similar to the worked examples (Stigler & Hiebert, 1999). The problem that faces the United States is as follows: How can U.S. preservice and in-service teachers make dramatic and fundamental changes in classroom practice that lead to students' acquiring a deeper understanding of mathematics and improving their achievement in mathematics?

Improving mathematical achievement involves changing both the instructional practices of teachers and the learning environment for students. According to the NCTM standards (NCTM, 2000), "mathematical expertise is found not only in factual knowledge, but also in the strategic decisions the student makes while solving problems and in their ability to communicate ideas about the problems with others" (Hiebert, 2003, p. 13).

Purpose of the Study

The purpose of the study was to determine the effectiveness of one professional development intervention, the MMI, as implemented in three high poverty (or low-socioeconomic) schools in a single school district. Changes in student achievement over a 3-year period were examined by using NCE (normal curve equivalency) scores from the SAT 10 Procedures and Problem Solving subtests and scale scores from the ARMT Math subtest. The independent variables used in the study were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

Null Hypotheses

The following null hypotheses were used in the present study:

1. There will be no significant interaction in mathematics achievement as measured by NCE scores on the Procedures subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose

teachers had participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

2. There will be no significant interaction in mathematics achievement as measured by NCE scores on the Problem Solving subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.
3. There will be no significant interaction in mathematics achievement as measured by scale scores on the ARMT Math subtest over a 2-year period for students in three high poverty schools from a single school system whose teachers participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

Significance of the Study

Although many professional development programs have been formulated in an attempt to improve mathematics skills of high poverty students, these studies have resulted in equivocal findings regarding success of these reform initiatives. Because the MMI has been demonstrated to be successful in improving student achievement of this targeted group, this study tested the effectiveness of the MMI in an additional location with a similar group of students (Van Hanegan, Pruet, & Bamberger, 2004).

It is anticipated that findings from this study will be useful to teachers, principals, curriculum specialists, mathematics specialists, professional development coordinators, and superintendents in the selection and implementation of nontraditional programs for teaching mathematics. If the MMI provides instruction that improves learning outcomes for these students, then other school systems may want to use a similar method for teaching mathematics that is evidence based in multiple settings.

Definition of Terms

ARMT is the Alabama Reading and Mathematics Test, which is based on the Alabama State Courses of Study and is a criterion-referenced test. The first full administration of this test was in spring 2005.

Constructivist learning is "an approach to teaching that gives learners the opportunity for concrete, contextually meaningful experience through which they can search for patterns, raise their own questions, and construct their own models, concepts, and strategies" (Fosnot, 1996, p. ix). The goal of constructivism in mathematics is to encourage students to invent various methods of problem solving and to make sense of the answers (Kamii, 1989).

Independent variables are variables that are thought to influence, affect, or cause outcomes (Creswell, 2003). The present study uses six independent variables: (a) ethnicity, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended. Definitions for the independent variables used in the study are as follow: *Ethnicity* is a grouping of people based on national origin (Gollnick & Chinn, 2002). *Gender* refers to the status of an individual as male or female and related issues (The Diversity Connection, n.d). According to the U.S. Bureau of the Census (as cited in Gollnick & Chinn), *socioeconomic status* is a criterion that measures an individual's economic condition. Individuals with *limited English proficiency* (LEP) are ones who are "non-English speakers or English language learners" (Gollnick & Chinn, p. 259). *Level of special education* refers to the need for "specially designed instruction, at no cost to parents, to meet the unique needs of a child with a disability" (U.S. Department of Education, n.d., para. 1).

Manipulative is a wide variety of physical materials and supplies that students use to foster the learning of abstract ideas in math. Examples include tools, models, blocks, tiles, and other objects (Catholic School Department, n.d.).

Mobile Math Initiative (MMI) is a year-round intensive mathematics professional development program for teachers and administrators. It operates on the premise that teacher quality directly impacts student achievement. The MMI follows the recommendations set forth by the NCTM (2000) Principles and Standards for School Mathematics.

National Assessment of Education Progress (NAEP; U.S. Department of Education, 2003) is an assessment of American students' knowledge and ability in Grades 4, 8 and 12 that is made on a continuing basis. NAEP, often called "The Nation's Report Card," has been administered regularly since 1969.

National Center for Education Statistics (NCES) is mandated by the U.S. Congress to

> collect, collate, analyze, and report full and complete statistics on the condition of education in the United States and other Nations; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; and assist state and local educational agencies in improving their statistical systems. (U.S. Department of Education, 1998, para. 6)

No Child Left Behind (NCLB) Legislation is the NCLB Act of 2001. This legisla-

tion

> is the reauthorization of a number of federal programs that strive to improve the performance of America's primary and secondary schools by increasing the standards of accountability for states, school districts, and schools, as well as providing parents more flexibility in choosing which schools their children will attend. (NCLB Act, p. 1)

PISA, the Program for International Student Assessment, was first implemented in 2000 by the Organization for Economic Cooperation and Development (OECD; 2003), an intergovernmental organization that is made of industrialized countries. In PISA 2003, mathematics literacy and problem solving were the areas in which data were gathered.

Principals and Standards for School Mathematics is a 2000 publication by the NCTM that provided guidelines calling for "all students to engage in more challenging mathematics" (p. 1).

Developmental education class (college) is a course that focuses on acquiring knowledge at precollege levels. These courses do not carry college credit (State University of New York, 1999).

Staff development is professional development that focuses on attitudes, skills and knowledge that teachers and administrators, as well as other school employees, need to be able to teach in such as way that all students can learn and achieve at high levels. Effective staff development includes training programs that are ongoing and of high quality with the provision for support and follow up (National Staff Development Council, 2001)

TIMSS 1995 is the Third International Mathematics and Science Study, a video study that documented typical teaching in Germany, Japan, and the United States. The video captured classroom processes used by national samples of eighth-grade mathematics teachers.

TIMSS-R 1999 is the Third International Mathematics and Science Study that expanded the first study by extending the video study to include typical teaching in the United States, Australia, the Czech Republic, Hong Kong, Japan, the Netherlands, and Switzerland.

Assumptions

This study involved several assumptions. These assumptions are as follows: (a) The National Mathematics Standards and the Alabama Mathematics Course of Study are valid guides for math instruction, (b) teachers trained in the use of the MMI are implementing MMI strategies daily in their classrooms as prescribed in the professional development section of the initiative, and (c) students who took the assessments demonstrated their best efforts.

Limitations

A limitation of the study was found with the ARMT Math subtest scores. This test is not administered in the second grade of any school; therefore, it was impossible to gather baseline data before these students entered the third grade. At that point, teachers had already been trained in the MMI, and this training further prevented the collecting of baseline data. Because baseline data were unobtainable for the ARMT Math subtest, it is not possible to detect the initial response of students to the first year of intervention.

Organization of the Study

Information about this study is reported in five chapters. Chapter 1 provides an introduction to the study, a statement of the problem, the significance of the problem, the purpose of the study, the research questions, the definitions of operational terms, the limitations faced in the research, and the assumptions. Chapter 2 is a review of the literature and includes national and international studies, U.S. efforts to improve mathematics achievement, exemplary and promising mathematics programs, and an overview of the MMI. Chapter 3 describes the methodological framework and procedures utilized in the study, including philosophical assumptions, design elements, data collection, data analysis, and ethical considerations. Methods of data collection are organized by topics. Chapter 4 reveals the findings of the research and includes the presentation and analysis of the data. The summary, findings, conclusions, implications, and recommendations for further research are discussed in chapter 5.

Summary

A deep understanding of mathematics with well-developed quantitative literacy skills is critical for 21st-century students. Given the relatively low international rankings of U.S. students in mathematics despite the development of NCTM standards and numerous attempts at reform, the ways in which students can most effectively learn to be mathematical thinkers and quantitatively literate citizens are not clear. It has been suggested that immediate reform efforts offer teachers the opportunity to transform their learning about mathematics and will lead to effective teaching practices that prepare students to be quantitatively literate citizens of the world.

CHAPTER 2

REVIEW OF LITERATURE

In this chapter, the review of literature is focused on data that indicate a continuing need for mathematics reform and on the steps that are being taken in the United States, as well as in the state of Alabama, to improve student achievement in mathematics. Literature related to four broad topics relevant to the study is reviewed. These topics are (a) mathematics reform efforts, which include national and international studies; (b) U.S. efforts to improve mathematics achievement; (c) exemplary and promising programs; and (d) an overview of the MMI.

Because college remediation costs the United States between $1 billion and $2 billion per year (Phipps, 1998), there is a growing concern about why students are not learning the prescribed curriculum in the primary and secondary grades. When attempting to think through this problem, educators will find it helpful to "begin with the end in mind" (Covey, 1989, p. 98). For example, investigators studying mathematics teaching and learning in Grades K-12 should know the skills needed in colleges and the workplace. *Greater Expectations: A New Vision for Learning as a Nation Goes to College* is a report from a consortium of college campuses involved in innovative 21st-century teaching (Association of American Colleges and Universities [AAC&U], 2002). The report states that the 21st-century student must

> become an intentional learner which develops self-awareness about the reason for study, the learning process itself, and how education is used. Intentional learners are integrative thinkers who can see connections in

> seemingly disparate information and draw on a wide range of knowledge to make decisions. They adapt the skills learned in one situation to new problems encountered in another—in a classroom, the workplace, their communities, and their personal lives. As a result, intentional learners succeed even when instability is the only constant. (AAC&U, p. 21)

In a recent Microsoft presentation, the facilitator, Sheryl Nussbaum-Beach (2005), mentioned the top four characteristics of a person whom the company would consider for employment. The characteristics were having the ability to effectively use thinking and problem-solving skills, possessing interpersonal skills, having communication skills, and being self-directed. Are U.S. primary and secondary schools developing these characteristics in students?

Recently, NCES of the U.S. Department of Education (2002) released the results of a study based on questions asked of 12th graders in the year 2000 about their perceptions of their educational experience. The same questions were asked in 1983, 1990, 1995, and 2000. In 1983, when asked to respond to the statement, "school work is often or always meaningful," 40% of students surveyed gave an affirmative answer; however, in 2000, only 28% answered "yes." When asked in 1983 whether school courses were "quite or very interesting," about 35% answered "yes"; in 2000, 21% answered the same question with a positive response. Last, students were asked whether school learning will be "quite or very important in later life." In 1983, 50% answered positively; however, in 2000, 39% answered in the affirmative (U.S. Department of Education, 2002).

Why do students perceive that there is a decline in the importance and meaningfulness of their work in school? Perhaps there is a disconnect between skills needed for 21st-century learners and the traditional teaching methods used in today's classroom.

Mathematics Reform Efforts

National Mathematics Studies

When pondering the need for educational reform, one finds a powerful warning in the release on April 26, 1983, of the government report *A Nation at Risk*. This disturbing report, which detailed the results of an 18-month study, was issued by a commission appointed by President Ronald Reagan and chronicled the poor condition of education in the United States. Included in the report was a statement that "the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a nation and a people" (*A Nation at Risk*, para. 1). Perhaps this document, which prefaced at least two educational initiatives launched by U.S. presidents, was the precursor to the NCLB Act of 2001. As Chester Finn (as cited in Coeyman, 2003), chairman of the Koret Task Force and a fellow at the Hoover Institute, said, "The report made a lasting contribution by changing national conversation about education" (p. 13).

A Nation at Risk (1983) contained discussions of some of the core beliefs of educational reform in this country. Although the document is more than 20 years old, the core values that it espoused continue to be important today. The values mentioned are as follows:

> 1. A high level of shared education is essential to a free, democratic society and to the fostering of a common culture, especially in a country that prides itself on pluralism and individual freedom. (*A Nation at Risk*, The Risk, para. 3)
>
> 2 All, regardless of race or class or economic status, are entitled to a fair chance and to the tools for developing their individual powers of mind and spirit to the utmost. This promise means that all children by virtue of their own efforts, competently guided, can hope to attain the mature and informed judgment needed to secure gainful employment, and to manage their own lives, thereby serving not only their own interests but also the process of society itself. (*A Nation at Risk*, The Risk, para. 5)

3. People are steadfast in their belief that education is the major foundation for the future strength of this country. They even considered education more important than developing the best industrial system or the strongest military force, perhaps because they understood education as a cornerstone of both. (*A Nation at Risk*, The Public's Commitment, para. 2)

In Marjorie Coeyman's (2003) article, "Twenty Years After 'A Nation at Risk'," published in the *Christian Science Monitor,* the author reminded Americans that, in comparison with people in other countries, U.S. citizens were falling behind in academics; that their standardized test scores were dropping; and that their schools were suffering from an abundance of low expectations. She went on to state that the document "brought us together to look at education with a sense of urgency" (Coeyman, p. 13).

The discussion of *A Nation at Risk* (1983) now serves as a springboard for the exploration of assessments, standards, and initiatives that relate to the teaching and learning of mathematics in this country. It is by analyzing data that one will find the road map to true educational reform.

National Assessment of Education Progress

Since 1969, NAEP (U.S. Department of Education, 2003) has been the only national assessment of what American students know and can do academically in the 4th, 8th, and 12th grades. The test is optional for 12th grade students. Mandated by the U.S. Congress, the assessment is administered by the NCES, which is part of the U.S. Department of Education. NAEP assesses various academic subjects, including mathematics, reading, writing, science, geography, U.S. history, civics, and the arts. The two major goals of the NAEP assessment are (a) to compare student achievement in states and jurisdictions and (b) to track changes in achievement over time. A sample of students is selected from public and nonpublic schools in each state. In the national sample, there are 10,000-20,000 students assessed from approximately 100 schools. In the state sample, there are about 3,000 students sampled per jurisdiction.

As defined by the NAEP, the three NAEP achievement levels in the mathematics framework are (a) basic, (b) proficient, and (c) advanced (U.S. Department of Education, 2003). *Basic* denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade level. *Proficient* represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter. The *advanced* achievement level represents superior performance.

A review of recent NAEP reports indicates that a majority of students are performing at a basic level. Although there has been a rise in the NAEP mathematics scores since 1990, there are still too few students reaching the proficiency level in mathematics. In 1990, 50% of fourth graders and 52% of eighth graders scored at or above the basic level as compared to 80% and 69%, respectively, in 2005. The proficiency level for fourth graders was 30% in 1990 and increased to 32% in 2005. The proficiency level for eighth graders went from 29% in 1990 to 30% in 2005. Although there has been some growth, the fact that less than one third of U.S. students are considered proficient in mathematics is evidence that there are major gaps in the teaching and learning of mathematics in elementary and secondary schools.

International Mathematics Studies

Program for International Student Assessment

PISA (OECD, 2003), which began in 2000, is administered every 3 years to assess the capabilities of 15-year-old students in reading literacy, *mathematics literacy*, and science literacy (Lemke et al., 2004). PISA defines mathematics literacy as “an individual’s capacity to identify and understand the role that mathematics plays in the world, to make well founded judgments in ways that meet the needs of that individual’s life as a constructive, concerned, and reflective citizen” (OECD, 2003, p. 24). The goal of PISA is to answer the question, “What knowledge and skills do students have at age 15?” This age was selected because it is within the age range for compulsory education. PISA is not based on grade level because 15-year-old students are in several grade levels and participate in many varied classes.

The purpose of PISA is to examine students’ abilities to apply a wide range of knowledge and skills to a variety of real-life problems such as personal, educational/occupational, public, or scientific (OECD, 2003). The assessment report is coordinated by the OECD, which is an intergovernmental organization made of 30 highly industrialized countries.

The 2003 PISA (OECD, 2003) focused on mathematics; the subareas of the mathematics literacy assessment were space and shape, change and relationships, quantity, and uncertainty. In comparison with other assessments, PISA 2003 used fewer multiple choice items than either NAEP (U.S. Department of Education, 2003) or TIMSS (Stigler & Hiebert, 1997) did; however, it focused more on data that used charts and graphs than the other assessments did. PISA 2003 was a 2-hr written assessment given to nationally representative samples. All 30 OECD countries participated in the assessment, as did 11 additional countries. In the United States, the 2003 PISA was administered to 5,456 public- and private-school students in 262 schools.

The U.S. overall average in mathematics literacy was below the OECD average. Twenty of the OECD countries and three non-OECD countries outscored the United States in mathematics. In each of the four subareas, the United States scored below the international average. Of the 38 countries that took the 2003 assessment, 24 outperformed the United States in the space-and-shape subtest, 21 outperformed the United States in the change-and-relationships subtest, 26 outperformed the United States in the quantity subtest, and 19 outranked the United States on the uncertainty scale (OECD, 2003).

Third International Mathematics and Science Study and Trends in International Mathematics and Science Study

The TIMSS was formerly known as the Third International Mathematics and Science Study (Stigler & Hiebert, 1997). TIMSS collects mathematics and science data at the fourth- and eighth-grade levels with the goal of providing data on trends in performance. Background information is gathered to address issues concerning the quality, quantity, and content of instruction. Data for this study were collected in 1995, 1999, and 2003 and will be collected in 2007. Approximately 50 countries participated in the study.

The 1995 and 1999 TIMSS videotape study (Stigler & Hiebert, 2000) confirmed that U.S. mathematics teachers who were observed emphasized computational procedures and spent little time in helping students develop concepts or in connecting procedures and concepts to show why the procedures work. The findings of the TIMSS study suggested that the traditional U.S. mathematics curriculum is not intellectually challenging but is repetitive and unfocused (Silver, 1998).

During the TIMSS 2003 study, 46 countries participated at the fourth- or eighth-grade level or both (Stigler & Hiebert, 2004). Conclusions were as follows:

1. There were no measurable changes detected between 1995 and 2003 in the average mathematics score of U.S. fourth graders. Moreover, the available data suggest that the performance of U.S. fourth graders in mathematics was lower in 2003 than in 1995 relative to the performance in those years of fourth graders in 14 other countries that participated in the studies.
2. U.S. fourth-grade girls showed no measurable changes between 1995 and 2003 in their average performance in mathematics. U.S. fourth-grade boys also showed no measurable change in their average mathematics performance.
3. U.S. Black fourth graders improved in mathematics between 1995 and 2003, whereas Hispanic fourth graders and White fourth graders showed no measurable changes in mathematics. As a result of the changes in the performance of Black fourth graders, the gap in achievement between Black and White fourth-grade students in the U.S. narrowed (Stigler & Hiebert, 2004).

U.S. Efforts to Improve Mathematics Achievement

National Council of Teachers of Mathematics Standards

The term *standards* entered the vocabulary of U.S. educators during the 1980s after the nation was declared at risk because of a mediocre educational system. In 1983, the National Commission on Excellence in Education called for schools and colleges to set more rigorous standards for students (Kilpatrick, 2001, p. 1).

The first group to develop mathematics standards was the NCTM. Presently having more than 100,000 members, NCTM is the world's largest mathematics organization. This professional organization offered Standards for Curriculum and Evaluation in 1989, Standards for Teaching in 1991, for Standards for Assessment in 1995, and Principles and Standards for School Mathematics in 2000.

To promote better student understanding of mathematics, NCTM (2000) developed standards that impact the way teachers teach mathematics. The standards suggest that the learning of facts and skills should also include conceptual understanding and a range of mathematical processes. The standards call for the ability to communicate mathematical ideas and to make strategic decisions when solving problems. The standards-based methods include building on students' entry-level skills, providing opportunities for intervention and practice, analyzing multiple methods, and giving students opportunities to discuss their explanations. The NCTM standards emphasize problem solving, making connections, reasoning, and communicating. These standards require teaching algebra, geometry, trigonometry, statistics, probability, discrete mathematics, and calculus. One of the goals of the standards-based curriculum is to change teacher practice by providing extensive, sustained, and focused professional development that gives teachers the opportunity to learn more effective pedagogy for teaching mathematics (NCTM).

When the first NCTM standards were published, some educators disagreed with the view of teaching and learning that called for students to work in small groups or use physical objects to explore mathematical ideas. During this period, skeptics advocated making no changes in mathematics teaching until the standards were validated by research. At this point, legislation began to be passed that called for materials, programs, and assessments to be validated by research (Kilpatrick, 2001, p. 1).

The National Science Foundation

President George H. W. Bush and all of the nation's governors attended an education summit in 1989. This summit served as the impetus to challenge the NSF to help U.S. schools to make changes in their mathematics curriculum on the basis of NCTM standards. The Education and Human Resource (EHR) division of the NSF offered a series of grants beginning in 1991 to encourage states to align their standards with NCTM standards.

In 1996, the NSF stated that effective standards-based education should incorporate the following beliefs:

1. All children can learn by using and manipulating scientific and mathematical ideas that are meaningful and relate to real-world situations and to real problems.
2. Mathematics and science are learned by doing rather than by using passive methods of learning such as watching a teacher work at the chalkboard. Inquiry-based learning and hands-on learning more effectively engage students than lectures do.
3. The use and manipulation of scientific and mathematical ideas benefit from a variety of contributing perspectives and are, therefore, enhanced by cooperative problem solving.
4. Technology can make learning easier, more comprehensive, and more lasting.

This view of learning is reflected in the professional standards of the NCTM, the American Association for the Advancement of Science (AAAS), and the National Research Council of the National Academy of Sciences (NRC).

Along with awarding grants in the 1990s, the NSF sponsored the creation of some commercial mathematics programs that were aligned with the NCTM standards. These programs are listed by elementary, middle, and high school levels, as follows: Elementary school programs include Everyday Mathematics; TERC's Investigations in Number, Data and Space; and Math Trailblazers. Middle school programs include Connected Math, Mathematics in Context, Mathscape: Seeing and Thinking Mathematically, MathThematics, and Pathways to Algebra and Geometry. High school programs include Contemporary Mathematics in Context, Interactive Mathematics Program, Math Connections: A Secondary Math Core Curriculum, Mathematics: Modeling Our World, and SIMMS Integrated Mathematics: A Modeling Approach Using Technology.

Promising and Exemplary Mathematics Programs

The U.S. Department of Education's Mathematics and Science Expert Panel (as cited in McGuire, 2000), released a report in 1999 that included descriptions of five exemplary mathematics programs and five promising programs. Discussions of two more promising programs were released in 2000 by the panel. To achieve exemplary status (the highest ranking), a program had to have received high marks on quality, usefulness to others, and educational significance and to have provided evidence that the program was effective in various sites and with various populations. Kent McGuire (2000), Assistant Education Secretary, said that these exemplary programs have met the highest standards set by our nation's leading mathematical experts and leaders:

> These programs work, and we will encourage teachers, administrators and policy makers to learn more about them as potential additions to their curriculum. The promising programs have great potential and strong but preliminary evidence that they too can serve our students well. (para. 3)

Promising Mathematics Programs

College Preparatory Mathematics (CPM) program. This program incorporates a 4-year secondary curriculum that integrates algebra and geometry with problem-solving skills. Six or seven core ideas are introduced, practiced, revisited, and mastered.

Cognitive Tutor Algebra. This program incorporates a first-year technology-based algebra course for secondary schools. Three days per week, students work on co-operative problem-solving activities; 2 days per week, they work independently in a laboratory. Students investigate real-world situations and use tools such as calculators and spreadsheets.

Connected Mathematics. This course of study is composed of a comprehensive problem-centered curriculum designed for Grades 6-8. Connections are made within mathematics, between various subjects and mathematics, and with the real world. Extensive problem sets are included to further understanding.

Core-Plus Mathematics Project. This program consists of a three-core course designed for all high school students. A fourth course is for the college-bound student. The curriculum focuses on mathematical modeling and the use of graphing calculators.

Interactive Mathematics Program (IMP). IMP is a 4-year, secondary, problem-based curriculum. Students are challenged to actively explore open-ended situations; investigate cases; look for patterns; and make, test, and prove conjectures.

Everyday Mathematics. This K-6 comprehensive mathematics program integrates mathematics with real-life situations. Features include problem solving, linking past experiences to concepts, sharing ideas, hands-on exploring, cooperative learning, and home/school partnering.

Mathland. Mathland is a K-6 program that uses problem solving to investigate and understand mathematics content. The program emphasizes skills such as problem solving; communicating; making connections; and using reasoning, estimation, statistics, measurement, probability, fractions, and decimals.

Middle School Mathematics through Application Project. In this Grades 6-8 technology-based curriculum, students use technology to analyze real-world problems, increase understanding of mathematics, and develop competence with standard symbolic notation for mathematical concepts and for organization and communication of ideas.

Number Power. This supplemental K-6 program uses a cooperative setting to develop number-sense and social-interaction skills. Students learn to do mental computation, estimate, analyze data, devise computation and problem-solving strategies, and compute accurately.

The University of Chicago School Mathematics Project (UCSMP). The project is a Grades 7-12 program composed of six courses that focus on real-world application and problem solving. The courses are Transition Mathematics, Geometry, Advanced Algebra, Functions/Statistics/Trigonometry, and Pre-calculus/Discrete Mathematics.

An Overview of the MMI

Thus far, the purpose of chapter 2 has been to examine national and international data with the intent of establishing a need for mathematics reform. Some general information was also presented on what is being done in the United States to increase student achievement in mathematics. Last, this chapter contains an overview of the MMI, a specific mathematics reform effort. This initiative served as the model for mathematics reform efforts taking place in single school district in Alabama (see Appendix A).

The MMI is herein described in conjunction with Alan Schoenfeld's (2002) requirements for sustained successful mathematics reform. Schoenfeld stated that the following four characteristics must be in place: (a) a high-quality curriculum; (b) a stable, knowledgeable, and professional teaching community; (c) high-quality assessment aligned with teaching goals; and (d) stability and the mechanisms that allow the curricula assessment and professional development to evolve effectively (p. 13). The first requirement for a high-quality curriculum was addressed in the MMI by basing all training on the Alabama Course of Study for Mathematics and on the NCTM (2000) curriculum and evaluation standards for school mathematics. The research-based strategies used in the training were modeled after Project Impact (Van Hanegan et al., 2004), which is based on the constructivist teaching of mathematics. During the training, there was an emphasis on content knowledge, pedagogy, and standards-based curricula.

A stable, knowledgeable, and professional teaching community is Schoenfeld's (2002) second criterion. In the 2 years of the MMI's operation in the targeted school district, only one school had principal turnover; the new principal was trained in the MMI, a requirement for all MMI school principals. Additionally, this principal had taught in an MMI school previously. In the 1st year of the initiative, fewer than five teachers in the eight schools had not been trained. In the 2nd year of implementation, a makeup session allowed all of the teachers to be trained. Teachers attended at least four additional workshops and received weekly coaching throughout the school year. Outside MMI consultants also visited teachers' classrooms several times a year and gave them feedback. There was a follow-up 3-day advanced workshop that took place the summer after the 2-week training, and teachers attending the advanced workshop were paid a stipend and received materials to implement what they learned. A 2-day mastery workshop was added during the second summer after the 2-week training; teachers received additional training, materials, and a stipend.

Schoenfeld's (2002) third criterion is "high quality assessment that is aligned with curricular goals" (p. 13). The full implementation of the ARMT in spring 2005 meant that students were assessed with a test that was more closely aligned with the curriculum that they were being taught. Students were also assessed with the SAT 10. The school district has created pacing guides with short monthly assessments to enable teachers to intermittently monitor students' progress in mathematics. The MMI also provides a quarterly problem for students; students solve the problem and explain the logic by which they arrived at the solution. Teachers have grade-level meetings to discuss the thinking that students used to answer these questions.

The fourth criterion is stability and mechanisms for the evolution of curricula, assessment, and professional development (Schoenfeld, 2002, p. 13). Some of the ways in which the MMI fulfills this condition within the system include system mathematics coaches who visit classes weekly, the Teacher Leadership Program, the Implementation Committee, Professional Math Learning Teams, outside-the-system coaches, grade-level meetings dedicated to mathematics, after-school workshops, during-the-workday workshops, ongoing principal training, summer training, pacing guides with lesson plans, Calendar Math training, and materials with which to implement what is learned in workshops.

Initially, the MMI focused on high-poverty/high-minority schools. Thirty-one percent of U.S. students entering college are considered to be minority, and 41% of students in the targeted school district are considered minority. Because these percentages represent a large portion of the student population, strategies for teaching minority students are herein explored by focusing on five areas that research indicates can affect mathematics achievement in minority students.

First, minority students are more likely to be taught with the use of low standards of performance. According to the ASCD Advisory Panel on Improving Student Achievement (Cole, 1995), low expectations are a barrier that interferes with effective instruction and student learning. Research shows that high expectations are necessary for student success. Often, minority students are placed in low-level mathematics courses and follow that track throughout their educational experience; however, all students, including minority students, need to be nurtured to develop positive attitudes about mathematics to increase the likelihood that they enroll in higher level mathematics courses commensurate with mathematical skills and abilities.

Second, the curriculum should meet the needs of the student and promote increased student achievement. Students learn mathematics best by "doing" or constructing the knowledge themselves rather than by copying how someone else does it. The goal of constructivism in mathematics is encourage students to invent various methods of problem solving and to make sense of the answers (Kamii, 1989).

> Traditionally mathematics has been taught in our schools as if it were a dead language. It was something that past, mostly dead, mathematicians had created—something that needed to be learned, practiced, and applied. When the definition of mathematics shifts toward "the activity of mathemitizing [constructing meaning for] one's lived world," the constructed nature of the discipline and its connection to problem solving become clear. When we define mathematics in this way, and teach accordingly, children will rise to the challenge. (Fosnot & Dolk, 2002, p. 18)

When minority students are allowed to apply their own reasoning to solving mathematical problems instead of being taught to simply memorize procedures, they increase their capacity for true mathematical understanding. For example, Ruby Payne (2001) suggests that the technology teacher might design a project that would require the student to use algebra to design a structure. When students are taught mathematics for understanding, they become better problem solvers and independent thinkers.

Third, every minority student should have a highly skilled teacher who is teaching in his or her area of expertise.

> Poor and minority children don't underachieve in school just because they often enter behind; but, also because the schools that are supposed to serve them actually *shortchange* them in the one resource they most need to reach their potential—high-quality teachers. Research has shown that when it comes to the distribution of the best teachers, poor and minority students do not get their fair share. (Peske & Haycock, 2006, p. 1)

Teachers must be trained to implement the best practices in the classroom. The NCTM (2000) developed national standards for the teaching of mathematics that moves away from rote memorization and toward a focus on understanding and reasoning. The

NCTM standards stress that (a) teachers should ensure that mathematics is interconnected and that several areas of mathematics are taught at every grade level (e.g., students should begin learning geometry in kindergarten and should study it in a more complex fashion as they progress through the grades); (b) teachers should teach mathematical problems in context and ensure that they reflect the real world of the students; (c) teachers should listen more and lecture less and should ask questions that both help students clarify their thinking and entice them to go to a deeper level of understanding; (d) teachers should help students to learn articulation skills and to become adept at explaining their thinking; and (e) teachers should help students to become familiar with charts, tables, and graphs and to use them to test, explore, and describe problems (NCTM).

Fourth, the school climate should honor diversity. All individuals in the school should acknowledge and value the race and ethnicity of all students, faculty, and staff by embracing differences, celebrating them, and incorporating them into the day-to-day happenings at the school. Even if the school is not highly diverse ethnically, the multiethnic nature of the United States should be reflected throughout the school—in the curriculum, on bulletin boards, in books, and in other resources used by the school. Teachers who pretend not to notice students' racial differences in fact do not notice the students at all; such teachers are limited in meeting the educational needs of their students (Ladson-Billings, 1994). Educators should develop the knowledge and skills needed to work in a diverse society. According to Gollnick & Chinn (2002), "social justice and equality for all people should be of paramount importance in the design and delivery of curricula" (p. 30). Valuing a student's culture facilitates learning and helps the student feel connected to the school culture. Students who feel that they are a part of a safe environment are more likely to take greater risks by taking more challenging classes and to excel in the classes they take.

Fifth, schools should provide numerous opportunities for parents and the community to become involved in the students' acquisition of mathematic skills. Parent training is invaluable, particularly in low socioeconomic communities (Payne, 2001). Family mathematics nights give parents strategies for working with their children on mathematics. Parents need to know that mathematics instruction has changed dramatically since they were children. Parents can be retrained to understand that teaching mathematics for understanding is different from offering strings of procedures to be memorized. Additionally, a productive community can participate in numerous opportunities designed to allow industry, businesses, civic clubs, Parent-Teacher Associations, and concerned individuals to tutor students or provide funding for math manipulatives and other materials needed for student learning. Effective mathematics education offers hope for a bright future for all students. No student should be left behind in the effort to teach mathematics for understanding.

CHAPTER 3

METHODOLOGY

In this study, a quasi-experimental research design was used to compare changes in students' mathematics achievement scores over time. The study design provides for comparisons across testing occasions, by using students as their own controls. Null hypotheses were as follows:

1. There will be no significant interaction in mathematics achievement as measured by NCE scores on the Procedures subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.
2. There will be no significant interaction in mathematics achievement as measured by NCE scores on the Problem Solving subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

3. There will be no significant interaction in mathematics achievement as measured by scale scores on the ARMT Math subtest over a 2-year period for students in three high poverty schools from a single school system whose teachers participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

The study is a longitudinal analysis of student achievement in a cohort of students before and after their teachers participated in professional development in the use of the MMI strategies over the course of 3 years. The main objective of the study was to examine changes in student achievement after implementation of MMI strategies.

The first year of student data were collected from the 2003-2004 school year and were used as baseline data for students whose teachers had not participated in MMI training. Teachers of these students received the MMI training during the summer before the MMI implementation in 2003-2004 and 2004-2005, when the students were in the third and fourth grades. In addition to analysis of changes in student achievement relative to teacher MMI training, an examination of changes in student achievement was conducted by using the independent variables of student ethnicity, gender, socioeconomic status, level of English proficiency, level of special education needed, and the specific school the student attends. Table 1 presents an overview of the MMI training schedule for teachers of study participants.

Table 1

MMI Teacher Training Schedule

2003-2004	2004-2005	2005-2006
Second grade (baseline): teachers have not been MMI trained	Third grade: first year for student to have an MMI-trained teacher	Fourth grade: second year for student to have an MMI-trained teacher
Priming for one teacher in each grade in kindergarten through fifth grade	Kindergarten and third-grade teachers trained summer 2004	First- and fourth-grade teachers trained summer 2005

Intervention

The professional development, or intervention, for teachers, mathematics coaches, and principals followed the curriculum specified in the MMI. The professional development incorporated best practices and provided instruction specific to each grade level. This process included an introduction to priming in the basic philosophy of the MMI during the academic year before the summer when the teachers were scheduled to participate in professional development. During the first summer of the first year of the MMI implementation, kindergarten and third-grade teachers were trained. The rationale for providing training to these particular grade levels was twofold. Training kindergarten teachers would result in students' benefiting from having their initial mathematics experiences with MMI-trained teachers. Training third-grade teachers would result in the school system's benefiting by being able to use standardized testing data (normally collected in the third grade) as a measure of student achievement with which to evaluate effectiveness of the MMI first-year implementation. The training schedule for subsequent years was for first- and fourth-grade teachers and second- and fifth-grade teachers, respectively. During the 2-week training, teachers were exposed to mathematics content and strategies based on the unique MMI professional development model described in Table 2.

Table 2

MMI Big Ideas

Curriculum	Instruction	Assessment
1. Problems/tasks that are relevant and engaging to students	1. Learning opportunities that allow students to make sense of mathematics, by building on prior knowledge	1. Instructional decisions made through ongoing and varied assessment
2. Coherent, relevant, and focused mathematics lessons	2. Clear, accurate, and precise teacher communication, both oral and written	2. Clear and precise oral and written communication by students
3. Connections between math topics and other content areas	3. Questioning strategies used to clarify, probe, and extend student thinking	
4. Technology used to enhance student learning	4. Challenging and engaging student learning opportunities through classroom environment and lessons	

Note: From "The Maysville to Mobile Mathematics Initiative from the Director's Perspective" by S. Pruet, 2005, In H. Bamberger (Ed.), *Mobile Math Initiative: Teaching Mathematics for Meaning and Mastery Summer Institute Participant's Notebook*, p. 1. Copyright 2005 by Mobile Area Education Foundation. Adapted with permission.

The following section describes the professional development provided to the teachers of student participants in the present study. Consultants who have been trained in the MMI strategies and who are employed by Mathworks under the direction of Dr. Honi Bamberger conduct the professional development workshops. Mathematics coaches and experienced MMI teachers co-present with the Mathworks presenters. Included in the workshops are opportunities for participants to individually practice their newly learned skills by teaching students enrolled in summer school. Feedback is provided to each teacher by a peer observer for five mornings during the midpoint of this 2-week intervention. Each day for 5 days, the observer facilitates a reflective conversation with the teacher. In addition, principals and teachers have faculty planning time during the workshop to restructure their curriculum to include the MMI strategies.

Four unique strategies that are daily routines for MMI-trained teachers and their students are Daily Data activities, Early Bird math activities, and THINK-PAIR-SHARE and other management activities.

1. Daily Data activities that include Venn diagrams and graphs are used to engage students in activities that make connections between real life and mathematics. Conversations about data collection provide students with numerous opportunities to explain their mathematical thinking through speaking and writing (Bamberger, 2005, p. 21).
2. Early Bird math activities are short, 10-min, independent, hands-on activities that reinforce math skills and concepts. Follow-up discussions about Early Bird activities strengthen vocabulary and language for students (Bamberger, 2005, p. 21).
3. THINK-PAIR-SHARE, a cooperative learning strategy, is a way to manage discussions with students and provide them with thinking time. Signal THINK by touching the pointer finger on the temple of your forehead to indicate to students to spend about 10-15 seconds thinking. Then signal PAIR (crossing the middle finger over the pointer finger) for students to talk quietly with someone near them about what they are thinking for 30-45 seconds. Signal SHARE (outstretched hand, palm up) to indicate it is time to share out loud either what they are thinking or something they heard. While students are SHARING, there are several additional management signals that ensure a productive discussion. The first signal is ME TOO, and is indicated by students pointing back and forth toward their chest with the thumb. This lets the teacher know they were thinking the same thing as the student sharing. The second signal is POINT OF INTEREST. Students use this signal by raising

their pointer finger in the air. It means they would like to add something or disagree with the person sharing. (Bamberger, 2005, p. 21)

4. "Equity sticks" are an additional strategy for managing any discussion with students. These are easily made by writing students' names on craft sticks. Store the sticks in a can. During a discussion, randomly choose sticks to call on students. (Bamberger, 2005, p. 21)

Specific strategies expected of teachers in the MMI workshops are congruent with best practices in mathematics (NCTM, 2000) and include the following strategies:

1. Plan instruction that is based on students' prior knowledge.
2. Facilitate additional learning opportunities by allowing adequate time for students to think and listen to each other.
3. Use questioning strategies that promote deep levels of understanding by the students. Teachers ask questions such as "How do you know that?" and "Will that work every time?" and "Did anyone get a different answer or work the problem another way?" In addition, students provide answers that justify math concepts.
4. Focus on students' thinking and problem-solving strategies instead of only on obtaining the correct answer.
5. Reinforce skills that enhance estimation and mental math.
6. Reinforce computational skills by connecting problem solving to real-world applications.
7. Integrate math throughout and across the curriculum.
8. Assess student knowledge and spontaneously modify instruction (NCTM, 2000).

In an MMI classroom, desks are grouped together for collaborative problem solving, and mathematic manipulatives are available for easy access. Calendar Math is evident in all classrooms. Student writing samples that show their mathematical thinking are displayed throughout the school building.

After the 2-week training, additional training opportunities are provided, as depicted in Table 3, and include a 3-day advanced-training workshop that teachers attend the summer after their 2-week training. Teachers are paid a stipend to attend workshops and are given several hundred dollars' worth of preselected grade-level-specific materials that are similar to the materials used in the 2-week summer institute to use in their classrooms; these materials include mathematics manipulatives.

Table 3

Professional Development Training Progression

Summer 1	Summer 2	Summer 3
Priming	Two-week summer institute	Three-day advanced institute

Levels of Personnel Involved With MMI Implementation

Although several unique features of the MMI are described in the literature, a specific feature of interest is that of the inclusion of math "coaches" hired to assist teachers in the classroom. Mathematics coaches focus on teachers who have been through the 2-week professional training. The coaches provide coaching and mentoring support to teachers, assist in planning, demonstrate lessons, co-teach, conduct observations, conduct monthly mathematics presentation to faculties, coordinate MMI plans with principals, coordinate follow-up classroom visits with consultants, participate in additional training, and keep records. Teachers from all eight MMI schools meet monthly and are provided with release time for planning sessions conducted within their schools. Additionally, the

MMI consultants provide quarterly visits to the classrooms of the MMI teacher participants to provide feedback.

Teacher Leadership Program

To support, enhance, and sustain the MMI innovation, teachers may apply to be selected to participate in the rigorous Teacher Leadership Program. The 3-year professional development program focuses on mathematics pedagogy, content, and leadership. In the 2nd and 3rd years of the program, participating teachers immerse themselves in a school-based leadership practicum.

Implementation Committee

To assist with implementing and sustaining the MMI, additional structures are necessary. One structure for both implementing the MMI with integrity and sustaining the MMI is the school's Implementation Committee. High implementers of the MMI are grade-level teachers who have integrated the MMI strategies into most of their teaching. These teachers are selected by the principals in the winter after their initial 10-day professional development. The primary functions of committee members are (a) to serve as a sounding board for mathematics coaches and district liaisons, (b) to be highly informed about the MMI, (c) to be advocates of the MMI, (d) to administer quarterly problems to students and to participate in discussion about student thinking related to these quarterly problems, and (e) to participate in professional learning teams (S. Pruet, personal communication, August 13, 2006).

Sample

The population for this study consisted of 89 students who were enrolled in three MMI schools in the selected school district; were second-grade students during the 2003-2004 school year; and remained in the same school through the 2005-2006 school year, when they were fourth-grade students. Enrollment in these three MMI schools includes Caucasian, African American, and Hispanic students. These schools were selected because they are composed of high poverty students, with over 80% receiving free or reduced-price lunch. Principals of these schools agreed to administer the SAT 10 to the school's second-grade students; the results of this testing provided pre-MMI baseline data. The demographics, including frequencies and sample size for ethnicity, gender, socioeconomic status, level of English proficiency, and level of special education needed, are included in Tables 4 through 7.

Table 4

Combined Demographics for Elementary Schools A, B, and C by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	16	18.0
Black	42	47.2
Hispanic	31	34.8
Gender		
Male	43	48.3
Female	46	51.7
Socioeconomic status		
Paid lunch	6	6.7
Free and reduced lunch	83	93.3
Level of English proficiency		
Non-limited English proficiency	60	67.4
Limited English proficiency	29	32.6
Level of special education needed		
Non-special education	78	87.6
Special education	11	12.4

Table 5

Elementary School A Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	5	16.1
Black	14	45.2
Hispanic	12	39.7
Gender		
Male	14	45.2
Female	17	54.8
Socioeconomic status		
Paid lunch	3	9.7
Free and reduced lunch	28	90.3
Level of English proficiency		
Non-limited English proficiency	19	61.3
Limited English proficiency	12	38.7
Level of special education needed		
Non-special education	24	77.4
Special education	7	22.6

Table 6

Elementary School B Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Demographic characteristics	*f*	% of sample
Ethnicity		
White	3	17.6
Black	13	76.5
Hispanic	1	5.9
Gender		
Male	7	41.2
Female	10	58.8
Socioeconomic status		
Paid lunch	2	11.8
Free and reduced lunch	15	88.2
Level of English proficiency		
Non-limited English proficiency	17	100.0
Limited English proficiency	0	0.0
Level of special education needed		
Non-special education	15	88.2
Special education	2	11.8

Table 7

Elementary School C Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Demographic characteristics	*f*	% of sample
Ethnicity		
White	8	19.5
Black	15	36.6
Hispanic	18	43.9
Gender		
Male	22	53.7
Female	19	46.3
Socioeconomic status		
Paid lunch	1	2.4
Free and reduced lunch	40	97.6
Level of English proficiency		
Non-limited English proficiency	24	58.5
Limited English proficiency	17	41.5
Level of special education needed		
Non-special education	39	95.1
Special education	2	4.9

As part of the University of Alabama at Birmingham's graduate research requirements, permission was obtained from the superintendent of the selected school district and from the university's Institutional Review Board for Human Use (Appendix B). All requirements for implementation of board regulations were followed during the data collection and data analysis phases of the study.

Data Collection Procedures

The data consisted of scale scores for the ARMT Math subtest from the 2004-2005 and 2005-2006 school years and of NCE scores for the SAT 10 Procedures and Problem Solving subtests from the 2003-2004, 2004-2005, and 2005-2006 school years. Table 8 provides an overview of the data collection plan. Student test scores for the SAT 10 and the ARMT from three high poverty schools that had participated in the MMI for 2 years were used in data analysis and were obtained from archival records in the school system's central office. Student names were not used in order to preserve the anonymity of students whose scores were used for statistical analysis. The central office of the school system provided access to the data, which were stored on a portable electronic storage device and locked in a metal file cabinet; student identity numbers were destroyed after the statistical analysis was conducted.

Table 8

Data Collection Schedule for Standardized Testing

2004	2005	2006
SAT 10 2nd grade	SAT 10 3rd grade	SAT 10 4th grade
	ARMT 3rd grade	ARMT 4th grade

As can be noted in Table 8, the process for data collection was systematic. Data were entered into Statistical Package for Social Sciences (SPSS; Version 15.0) for subsequent analysis. The data collected from three high poverty schools specifically included the following student test scores for the ARMT Math subtest: (a) spring 2005 ARMT Math subtest scale scores for third-grade students from three MMI schools in the targeted school district and (b) spring 2006 ARMT Math subtest scale scores for fourth-grade students from three MMI schools in the targeted school district. The data collected from three high poverty schools included the following students' test scores for the SAT 10 Procedures and Problem Solving subtests: (a) spring 2004 SAT 10 Procedures and Problem Solving subtest NCE scores for second-grade students from three MMI schools in the targeted school district, (b) spring 2005 SAT 10 Procedures and Problem Solving subtest NCE scores for third-grade students from three schools in the targeted school district; and (c) spring 2006 SAT 10 Procedures and Problem Solving subtest NCE scores from fourth-grade students in three schools in the targeted school district.

Data Analysis Plan

The data analysis plan addressed each of the three null hypotheses by using factorial ANOVA with repeated measures analysis procedures. In this study, students were treated as their own control. In addition, this longitudinal study provides data to describe changes in achievement over the 3-year period.

The interaction among achievement measures with the following independent variables was compared: (a) ethnic background (White, Black, Hispanic), (b) gender (male, female), (c) socioeconomic status (paid lunch, free and reduced lunch), (d) level of English proficiency (non-limited English proficient, limited English proficient), (e) level of special education needed (non-special education, special education), and (f) school student attended. These comparisons involve data about students who participated in the SAT 10 Procedures and Problem Solving subtests for the school years 2003-2004, 2004-2005, and 2005-2006 and in the ARMT Math subtest for the school years 2004-2005 and 2005-2006.

Instruments

Two instruments, the ARMT and the SAT 10, provided the data used to measure the dependent variable. Both of these tests are standardized, and the publishers of both assessments have reported on the psychometric properties taken to assure the validity and reliability of the testing instruments. These assessments are both reliable and valid for use in comparing changes in student achievement over time.

ARMT Validity and Reliability

Validity of the ARMT

The most basic element of test development and evaluation is validity. Validity "means that researchers can draw meaningful and justifiable inferences from scores about a sample or population" (Creswell, 2005, p. 600). The technical report (*Alabama Reading and Mathematics Test,* 2005) developed for use with the ARMT indicates that construct validity is the most important type of validity for use of this assessment. Construct validity of the ARMT was examined by using the intercorrelations of the specified domains, subdomains, and total scores provided in Table 9.

Table 9

Intercorrelations of Domains, Subdomains, and Total Scores for the ARMT Math Subtest for Grade 3

Mathematics 3	MT	PR	PS	NSO	PRA	GMY	MST	DSP
Math Total (MT)	1.00							
Procedures (PR)	.76	1.00						
Problem Solving (PS)	.99	.66	1.00					
Number Sense And Operations (NSO)	.85	.92	.78	1.00				
Algebra (PRA)	.70	.49	.70	.56	1.00			
Geometry (GMY)	.76	.47	.78	.54	.47	1.00		
Measurement (MST)	.62	.36	.64	.42	.34	.44	1.00	
Data Analysis & Probability (DSP)	.69	.41	.71	.48	.41	.49	.41	1.00

For Mathematics 3, the intercorrelation between Math Total and Problem Solving is .99, whereas the intercorrelation between Math Total and Procedures is .76. "The results imply that Mathematics Tests were mainly designed to measure Problem Solving rather than Procedures" (*Alabama Reading and Mathematics Test,* 2005, p. 60) for Grade 3 (this result was also found for Grades 5 and 7). On the basis of intercorrelation values, the researcher expected that mean scale scores for the ARMT Math subtest would be more closely aligned with mean NCE scores for the SAT 10 Problem Solving subtest than with mean NCE scores for the SAT 10 Procedures subtest.

Reliability of the ARMT

Test reliability indicates the precision and consistency of measurement. Reliability coefficients and standard error of measurement are two statistics that describe a test's reliability. The types of reliability that are considered for the ARMT are internal consistency and test-retest reliability (*Alabama Reading and Mathematics Test*, 2005, p. 46).

Interrater agreement on open-ended test items. The amount of interrater agreement was estimated by using a check score procedure. Data from the check score procedure were analyzed by using the percentage of perfect agreement between adjacent score categories. If the discrepancies between the first and second rating are within ±1 score, then there is interrater agreement. The data presented in Table 10 indicate that the ARMT is a valid measure for assessing student achievement.

Table 10

Interrater Agreement Coefficients for the ARMT Math Subtest for Grade 3

	Interrater agreement	
Subject and grade	% Perfect	+1 Adjacent
Math 3	81	98

SAT 10 Validity and Reliability

The SAT 10, developed by Harcourt, uses the definition of *validity* that is published in the current edition of the Standards for Educational and Psychological Testing; this definition states that validity is "the degree to which evidence and theory support the interpretations of test scores entailed in the use of tests [and] is, therefore, the most fundamental consideration in developing and evaluating tests" (*Stanford Spring Technical Data Report*, 2003, p. 9). The SAT 10 incorporates the 24 validity-related standards (1.1 through 1.24) from the same publication. The following sources of validity were used to evaluate test validity of the SAT 10: (a) test content, (b) test response processes, (c) test internal structure, (d) relationships to other variables, (e) convergent and discriminant analysis, (f) test criterion relationships, and (g) testing consequences. Examination of these data provided in the technical manual indicated that the evidence for reliability of the tests at all grades was high.

Test reliability is reflected "in evidence of test accuracy, precision, and consistency" (*Stanford Spring Technical Data Report*, 2003, p. 44). The reliability coefficient that demonstrates internal consistency emphasizes the consistency of test performance from item to item (see Table 11). This consistency is accomplished by

> subdividing a test into portions, typically halves, and correlating the scores from each portion. To overcome the possibility of non-equivalent portions the Kuder-Richardson Formula 20 (KR20) is used to generate the KR20 reliability coefficient. The KR20 reliability coefficients show that the Stanford 10 is reliable and based on a high degree of internal consistency. (*Stanford Spring Technical Data Report*, p. 45)

The *Stanford Spring Technical Data Report* offers further descriptions of the validity and reliability measures for the SAT 10.

Table 11

Reliability Coefficients for the SAT 10 Mathematics Subtests for Grade 3

Grade 3 subtest	Number of items	*N*	*M*	*SD*	*SEM*	KR-20
Problem Solving	30	2,170	20.1	6.2	2.26	.87
Procedures	20	2,177	12.2	4.8	1.84	.85

Summary

Chapter 3 included a description of the research design, intervention, sample, and data employed in the study. In addition, descriptions of the criteria used for selecting the sample and of the population demographics from which the samples were derived were provided. From the population of second-, third-, and fourth-grade students, 89 students with available scores for the three testing occasions were included in the study. The SAT

10 Procedures and Problem Solving subtests and the ARMT Math subtest provided the basis for the data that were gathered. In the study, the effects on student achievement that result when the student is taught by a trained MMI teacher were explored by examining the relationship of mathematics achievement scores to the independent variables of (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended; the three null hypotheses were examined by using a factorial ANOVA with repeated measures.

CHAPTER 4

FINDINGS

Introduction

The purpose of this chapter is to report findings for the three null hypotheses related to student achievement and the MMI. The assumptions for ANOVA with repeated measures have been discussed. Data from the 89 students enrolled in the three MMI schools were analyzed by using the SPSS (Version 15.0). Findings are reported for each of the three null hypotheses.

Data Verification

Accuracy of data entry was verified by a research methodology committee member and a graduate student to ensure the validity of data for use in analyses. Before being analyzed, the data set was screened for outliers as a further step to increase the accuracy of the results reported in this chapter. Inclusion of these scores in the data analysis resulted in a skewed distribution that yielded less-than-valid results. Scores of 10 were significantly different from the next highest score. The minimum score for the SAT 10 NCE scores was set at 10; therefore, any score less than 10 was considered an outlier and filtered out of the data set. According to Osborne and Overbay (2004), the most appropriate decision is to remove those outliers when such outliers occur. Seven NCE scores less than 10 were filtered from the SAT 10 Procedures 2004 subtest, and six NCE scores were filtered from the SAT 10 Problem Solving 2004 subtest. Four scores that were less than 540 in the ARMT Math subtest for 2005 were filtered, as well. Thus, the following scores were used in the analysis at baseline: 2004 SAT 10 Procedures subtest NCE scores (n = 82), 2004 SAT 10 Problem Solving subtest NCE scores (n = 83), and 2005 ARMT Math subtest scale scores (n = 85).

Procedures for Repeated Measures

The purpose of this study was to investigate the change in performance of students whose teachers were trained in the MMI. The research design provided for a longitudinal analysis in which students were used as their own controls. SAT 10 Procedures and Problem Solving scores over a period of 3 consecutive years were examined, as were the ARMT Math subtest scores for 2 consecutive years. In this study, the dependent variables were NCE scores for the SAT 10 Procedures and Problem Solving subtests and the scale scores for the ARMT Math subtest. Student achievement data in standardized form were used. Data were analyzed by using the following independent variables: (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended. An ANOVA procedure with repeated measures was chosen as the method of analysis because this method allows comparisons across testing occasions by using students as their own controls (Green, Salkind, & Akey, 1999). A power analysis was conducted for the samples, and there was found to be sufficient power (power = .80) to detect significant interactions for schools and testing occasions but not to detect significant interactions for the other independent variables. Thus, data for those variables were examined as the main effects of the model.

Assumptions of Repeated Measures

The following assumptions of repeated measures described by Mertler and Vannatta (2001) were considered in this study:

1. The observations within each sample must be randomly sampled and must be independent of one another. Because the design was a repeated measures design using subjects as their own control, meeting this assumption was not required.
2. The distributions of scores on the dependent variable must be normal in the populations from which the data were sampled. This assumption was met after the removal of outliers.
3. The assumption of sphericity, which is the most critical assumption in the multivariate model as related to homogeneity, was met. In the multivariate model (repeated measures ANOVA), Mauchly's test of sphericity was used in the analysis (p = .458 for Null Hypothesis 1; p = .097 for Null Hypothesis 2). Because $p > .05$ for Mauchly's test, sphericity was assumed for all hypotheses (Davis, 2002).

Demographic Characteristics of the Sample

A total of 89 elementary students in the targeted school district participated in the study. Of the 89 students, 31 (34.8%) attended School A, 17 (19.10%) attended School B, and 41 (46.1%) attended School C. Table 12 displays demographics for the three schools participating in the study, including frequencies and sample size for ethnicity, gender, level of English proficiency, and level of special education needed. Individual demographics for the three schools in the study are presented in Tables 13, 14, and 15.

Table 12

Combined Demographics for Elementary Schools A, B, and C by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	16	18.0
Black	42	47.2
Hispanic	31	34.8
Gender		
Male	43	48.3
Female	46	51.7
Socioeconomic status		
Paid lunch	6	6.7
Free and reduced lunch	83	93.3
Level of English proficiency		
Non-limited English proficiency	60	67.4
Limited English proficiency	29	32.6
Level of special education needed		
Non-special education	78	87.6
Special education	11	12.4

Table 13

Elementary School A Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	5	16.1
Black	14	45.2
Hispanic	12	39.7
Gender		
Male	14	45.2
Female	17	54.8
Socioeconomic status		
Paid lunch	3	9.7
Free and reduced lunch	28	90.3
Level of English proficiency		
Non-limited English proficiency	19	61.3
Limited English proficiency	12	38.7
Level of special education needed		
Non-special education	24	77.4
Special education	7	22.6

Table 14

Elementary School B Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	3	17.6
Black	13	76.5
Hispanic	1	5.9
Gender		
Male	7	41.2
Female	10	58.8
Socioeconomic status		
Paid lunch	2	11.8
Free and reduced lunch	15	88.2
Level of English proficiency		
Non-limited English proficiency	17	100.0
Limited English proficiency	0	0.0
Level of special education needed		
Non-special education	15	88.2
Special education	2	11.8

Table 15

Elementary School C Demographics by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	8	19.5
Black	15	36.6
Hispanic	18	43.9
Gender		
Male	22	53.7
Female	19	46.3
Socioeconomic status		
Paid lunch	1	2.4
Free and reduced lunch	40	97.6
Level of English proficiency		
Non-limited English proficiency	24	58.5
Limited English proficiency	17	41.5
Level of special education needed		
Non-special education	39	95.1
Special education	2	4.9

Findings of Statistical Analysis for Null Hypothesis 1

ANOVA with repeated measures was used to test Null Hypothesis 1: There will be no significant interaction in math achievement as measured by NCE scores on the Procedures subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The analysis was performed by using SPSS (Version 15.0). Descriptive statistics were computed to provide additional understanding of the data. As previously noted, the dependent variable, student achievement, and the independent variables, (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended, were examined. The dependent variable examined was NCE scores for the SAT 10 Procedures subtest from 2004-2006. Table 16 provides descriptive statistics for NCE scores for the SAT 10 Procedures subtest that were reported by school.

Table 16

Descriptive Statistics for NCE Scores for the SAT 10 Procedures Subtest for All Testing Occasions by School

	2004			2005			2006		
School	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
A	25	26.63	12.01	25	44.85	19.87	25	53.77	20.96
B	16	45.11	23.30	16	49.52	24.81	16	60.76	21.89
C	41	44.46	22.17	41	44.88	19.60	41	46.27	18.03

Statistically Significant Findings for Null Hypothesis 1

The null hypothesis of no significant interaction was rejected for the independent variable, school ($F = 7.78$, $df = 4$, $p < .001$). These findings showed statistical significance between time (testing occasions) and school. As shown in Table 17, Pillai's Trace value ($p < .001$) indicates the significance in the multivariate model. In addition, there was significant interaction between scores and time (testing occasions; $F = 25.57$, $df = 2$, $p < .001$). The interaction between time (testing occasions) and school is linear and accounts for approximately 29.8% of variability in the dependent variable, year. The interaction between scores and time (testing occasions) also is linear. Descriptive statistics for all independent variables for all testing occasions are presented in Table 18.

Table 17

Multivariate Tests for NCE Scores for the SAT 10 Procedures Subtest

Effect	Source	*F*	*df*	Significance	Partial η^2
Time	Pillai's Trace	25.57	2	<.001	.396
Time * School	Pillai's Trace	7.78	4	<.001	.165
Time	Linear	51.53	1	<.001	.395
Time * School	Linear	16.76	2	<.001	.298

Table 18

Descriptive Statistics for NCE Scores for the SAT 10 Procedures Subtest for Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed for All Testing Occasions

	2004			2005			2006		
Variable	*n*	*M*	*SD*	*N*	*M*	*SD*	*n*	*M*	*SD*
Ethnicity									
White	14	39.16	16.80	14	40.74	24.09	14	45.30	20.81
Black	38	39.69	22.61	38	44.06	20.19	38	54.39	20.92
Hispanic	30	38.93	22.21	30	50.31	19.15	30	50.42	19.16
Gender									
Male	42	39.58	21.58	42	47.06	21.27	42	49.02	20.36
Female	40	39.70	21.37	40	44.04	20.03	40	53.86	20.18
Socioeconomic status									
Paid lunch	76	39.00	21.65	76	45.26	20.35	76	50.91	20.49
Free and reduced lunch	6	41.12	18.61	6	52.77	24.37	6	57.40	18.03
Level of English proficiency									
Non-limited English proficiency	54	39.97	21.18	54	44.25	21.53	54	51.48	20.89
Limited English proficiency	28	37.58	21.97	28	48.91	18.68	28	51.19	19.47
Level of special education needed									
Non-special education	74	41.73	20.71	74	48.36	19.67	74	54.10	18.84
Special education	8	15.25	8.87	8	21.90	12.21	8	26.29	16.15

Line graphs were plotted to illustrate the trend of performance on the SAT 10 Procedures subtest across testing occasions. As shown in Figure 1, there was a constant and linear increase in mean scores for this subtest across testing occasions.

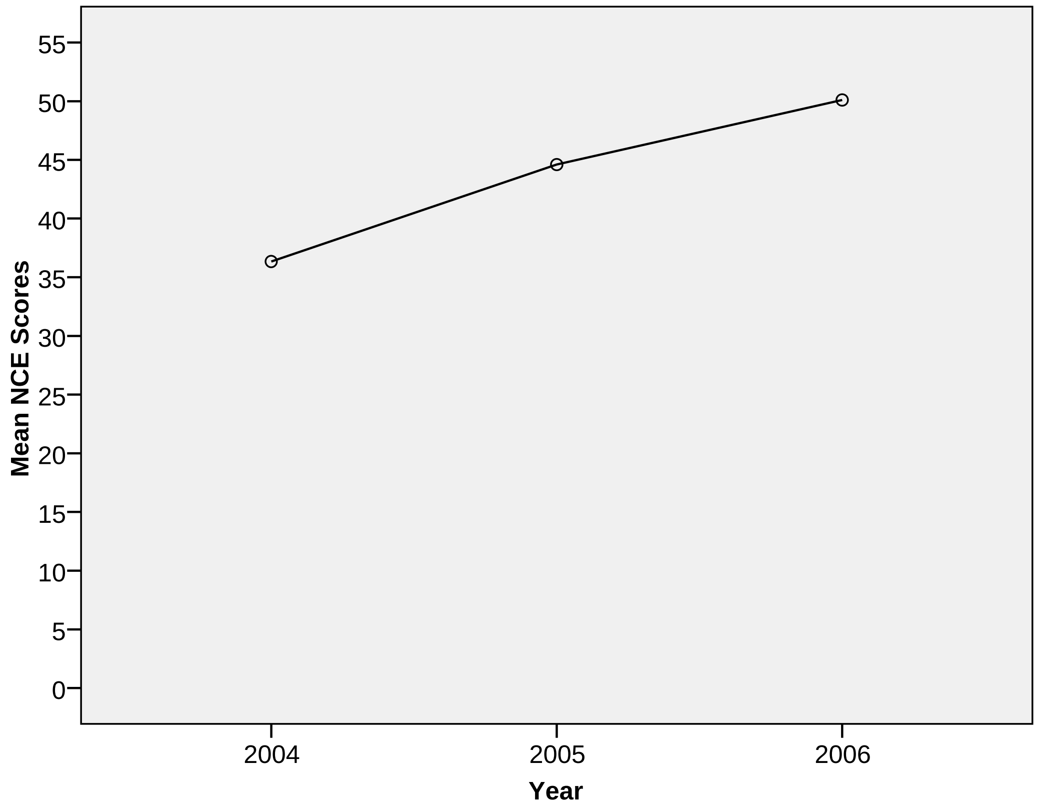

Figure 1. NCE scores for the SAT 10 Procedures subtest across testing occasions.

Although the findings showed statistical significance between time (testing occasions) and school, there was insufficient evidence to reject Null Hypothesis 1 for the other independent variables. This finding can be explained, in part, by the small sample sizes in many of the cells. The sample sizes were too small to detect a statistically significant interaction even if one existed. Thus, these variables were examined as main effects by using independent *t* tests, paired *t* tests, and one-way ANOVA. To determine whether the differences in means between categories of independent variables were statistically significant at baseline, at one year of intervention, and at 2 years of intervention, the researcher conducted independent *t* tests for gender, socioeconomic status, level of English proficiency, and level of special education needed, and one-way ANOVA for ethnicity. To determine whether the differences over time of all categories of independent variables were significant, paired *t* tests were conducted. Results of independent *t* tests are presented in Table 19, those for one-way ANOVA are given in Table 20, and those for paired *t* tests are displayed in Tables 21 and 22.

Table 19

Independent t Tests for NCE Scores for the SAT 10 Procedures Subtest for Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	Year	*t*	Significance (2-tailed)
Gender	2004	0.184	.854
	2005	0.573	.568
	2006	-1.080	.283
Socioeconomic status	2004	-0.233	.816
	2005	-0.862	.391
	2006	-0.752	.454
Level of English proficiency	2004	0.478	.634
	2005	-0.929	.355
	2006	0.061	.952
Level of special education needed	2004	3.566	.001
	2005	3.715	<.001
	2006	4.020	<.001

Table 20

One-way ANOVA for Ethnicity for NCE Scores for the SAT 10 Procedures Subtest

Year	*df*	*F*	Significance
2004	2, 79	0.028	.972
2005	2, 79	1.288	.281
2006	2, 79	1.082	.344

Table 21

Paired t Tests for NCE Scores for the SAT 10 Procedures Subtest Between 2004 and 2006 for School, Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Level of categorical pairing for 2004 and 2006	*t*	*df*	Significance
School			
School A	-7.662	24	<.001
School B	-3.904	15	.001
School C	-0.651	40	.519
Ethnicity			
White	-1.657	13	.121
Black	-4.427	37	<.001
Hispanic	-2.700	29	.009
Gender			
Male	-2.918	41	.006
Female	-4.807	39	<.001
Socioeconomic status			
Free and reduced lunch	-4.923	75	<.001
Paid lunch	-3.266	5	.022
Level of English proficiency			
Limited English proficiency	-3.314	27	<.001
Non-limited English proficiency	-4.212	53	<.001
Level of special education needed			
Special education	-2.171	7	.067
Non-special education	-5.020	73	<.001

Table 22

Paired t Tests for NCE Scores for the SAT 10 Procedures Subtest Between 2004 and 2005 and Between 2005 and 2006 for School

	2004-2005			2005-2006		
School	*t*	*df*	Significance	*t*	*df*	Significance
A	-6.249	24	<.001	-3.712	24	.001
B	-0.982	15	.341	-2.100	15	.053
C	-0.154	40	.879	-0.609	40	.546

As shown in Figure 2, Schools A and B experienced statistically significant gains in NCE scores for the SAT 10 Procedures subtest between 2004 and 2006; furthermore, School A experienced statistically significant gains at each testing occasion. School C experienced gains in SAT 10 Procedures subtest NCE scores that were not found to be statistically significant.

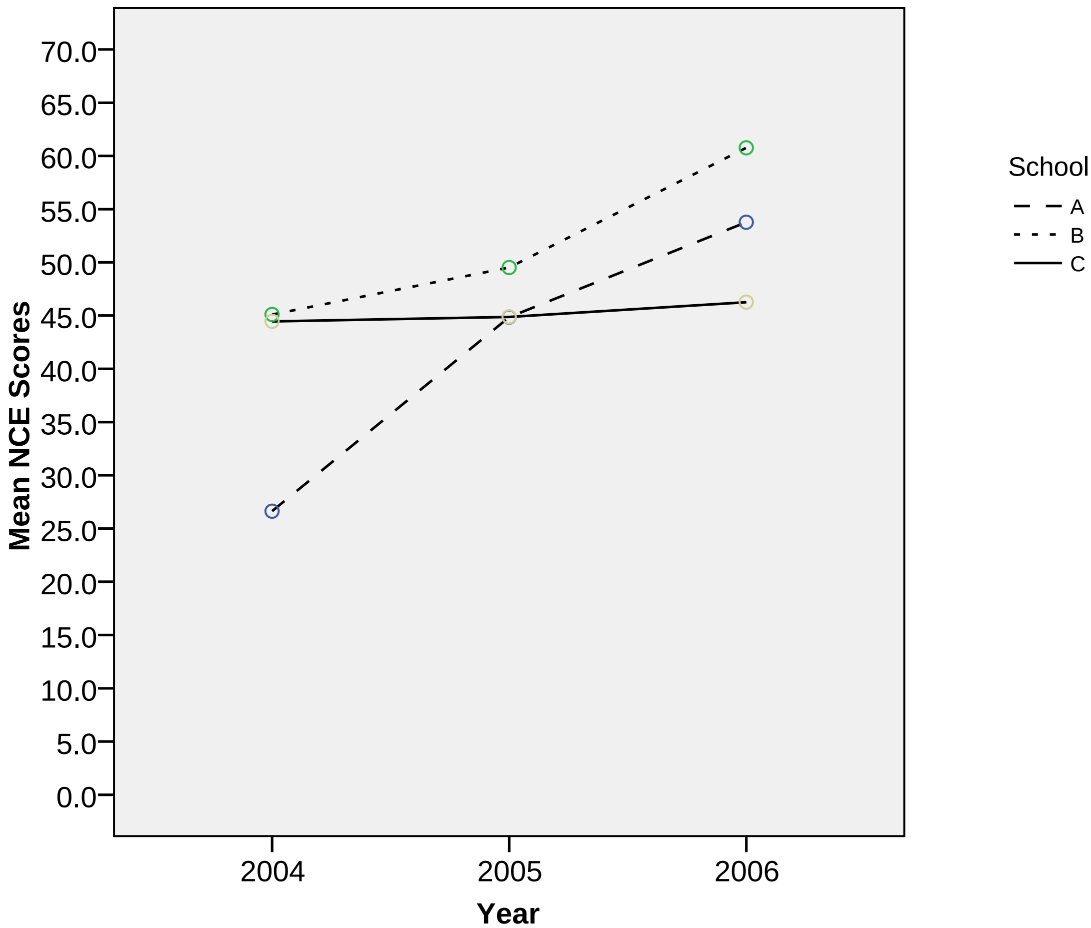

Figure 2. NCE scores for the SAT 10 Procedures subtest reported by school.

The following findings are reported by independent variable:

1. At baseline, mean NCE scores for the SAT 10 Procedures subtest for Black and Hispanic students were similar to mean NCE scores for White students (M = 39.69 for Black students; M = 39.93 for Hispanic students; M = 39.16 for White students); after 2 years of intervention by an MMI-trained teacher,

mean NCE scores for the SAT 10 Procedures subtest for Black and Hispanic students were higher than mean NCE scores for White students (*M* = 54.39 for Black students; *M* = 50.42 for Hispanic students; *M* = 45.30 for White students). On the basis of results of one-way ANOVA, the difference in mean NCE scores for the SAT 10 Procedures for ethnicity was not found to be statistically significant at baseline ($F = 0.028$, $df = 2, 79$, $p = .972$), or after 2 years of intervention ($F = 1.082$, $df = 2, 79$, $p = .344$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for Black students over the 3 years of the study showed a statistically significant gain ($t = -4.427$, $df = 37$, $p < .001$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for Hispanic students over the 3 years of the study showed a statistically significant gain ($t = -2.790$, $df = 29$, $p = .009$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for White students over the 3 years of the study failed to show a significant gain ($t = -1.657$, $df = 13$, $p = .121$).

2. At baseline, mean NCE scores for the SAT 10 Procedures subtest for female students were slightly higher than mean NCE scores for male students (*M* = 39.58 for male students, *M* = 39.70 for female students); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for female students were considerably higher than mean NCE scores for male students (*M* = 49.02 for male students, *M* = 53.86 for female students). On the basis of results of independent *t* tests, the difference in mean NCE scores for the SAT 10 Procedures subtest for gender was not found to be statistically significant at baseline ($t = 0.184$, $df = 80$, $p = .854$), or after 2 years of intervention ($t = -1.080$, $df =$

80, $p = .283$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for male students over the 3 years of the study showed a statistically significant gain ($t = -2.918$, $df = 41$, $p = .006$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for female students over the 3 years of the study showed a statistically significant gain ($t = -4.807$, $df = 39$, $p < .001$).

3. At baseline, mean NCE scores for the SAT 10 Procedures subtest for students who received subsidies for lunches were similar to mean NCE scores for students who paid for their lunches ($M = 39.00$ for students who received subsidies for lunches; $M = 41.12$ for students who paid for their lunches); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the Procedures subtest for students who paid for their lunches remained higher than mean NCE scores for students who received subsidies for their lunches ($M = 50.91$ for students who received subsidies for their lunches; $M = 57.40$ for students who paid for their lunches). On the basis of independent *t* tests, the difference in mean scores for the SAT 10 Procedures subtest for socioeconomic status was not found to be statistically significant at baseline ($t = -0.233$, $df = 80$, $p = .816$), or after 2 years of intervention ($t = -0.752$, $df = 80$, $p = .454$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for students who received subsidies for their lunches showed a statistically significant gain ($t = -4.923$, $df = 75$, $p < .001$) over the 3 years of the study. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for students who paid for their lunches showed a statistically significant gain ($t = -3.266$, $df = 5$, $p = .022$).

4. At baseline, mean NCE scores for the SAT 10 Procedures subtest for LEP students were lower than mean NCE scores for non-LEP students (*M* = 37.58 for LEP students; *M* = 39.97 for non-LEP students); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Procedures subtest for LEP students remained lower than mean NCE scores for non-LEP students (*M* = 51.19 for LEP students; *M* = 51.48 for non-LEP students). On the basis of results of independent *t* tests, the difference in mean NCE scores for the SAT 10 Procedures subtest for English proficiency was not found to be statistically significant at baseline ($t = 0.478$, $df = 80$, $p = .634$), or after 2 years of intervention ($t = 0.061$, $df = 80$, $p = .952$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for LEP students over the 3 years of the study showed a statistically significant gain ($t = -3.314$, $df = 27$, $p = .003$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for non-LEP students over the 3 years of the study showed a statistically significant gain ($t = -4.212$, $df = 53$, $p < .001$).
5. At baseline, mean NCE scores for the SAT 10 Procedures subtest for special education students were lower than mean NCE scores for non-special education students (*M* = 15.25 for special education students; *M* = 41.73 for non-special education students); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Procedures subtest for special education students continued to be lower than mean NCE scores for non-special education students (*M* = 26.29 for special education students; *M* = 54.10 for non-special education students). On the basis of results of independent *t* tests, the difference in mean NCE scores for the SAT 10 Procedures subtest for spe-

cial education was found to be statistically significant at baseline ($t = 3.566$, $df = 80$, $p = .001$), and after 2 years of intervention ($t = 4.020$, $df = 80$, $p < .001$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for special education students over the 3 years of the study failed to show a statistically significant gain ($t = -2.171$, $df = 7$, $p = .067$); however, a trend was noted toward statistical significance. (A trend toward statistical significance is noted when $p \leq 1.0$; G. Abbott, personal communication, May 11, 2007.) Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for non-special education students over the 3 years of the study showed a statistically significant gain ($t = -5.020$, $df = 73$, $p < .001$).

Conclusions for Null Hypothesis 1

For Null Hypothesis 1, the null hypothesis of no significant interaction was rejected for the independent variable, school ($F = 7.78$, $df = 4$, $p < .001$). These findings showed statistical significance between time (testing occasions) and school. In addition, there was significant interaction between scores and time (testing occasions; $F = 25.57$, $df = 2$, $p < .001$). There was insufficient evidence to reject Null Hypothesis 1 for the other independent variables. This finding can be explained in part by the small sample sizes in many of the cells. Independent *t* tests comparing scores for categories of independent variables at baseline and after 2 years of intervention by an MMI-trained teacher were only significant for special education ($p = .001$, $p < .001$, respectively). Paired *t* tests comparing mean NCE scores for the SAT 10 Procedures subtest for each category of independent variable at baseline and final testing occasions (2004 and 2006, respectively) were significant for Black, Hispanic, male, female, subsidized student, non-subsidized student, LEP, non-LEP, and non-special education students, and not significant for White and special education students.

Findings of Statistical Analysis for Null Hypothesis 2

ANOVA with repeated measures was used to test Null Hypothesis 2: There will be no significant interaction in mathematics achievement as measured by NCE scores on the Problem Solving subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The analysis using SPSS (Version 15.0) indicated that Null Hypothesis 2 was rejected for all schools across testing occasions. Descriptive statistics, displayed in Table 23, provide an additional examination of the data. As in Null Hypothesis 1, the independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

Table 23

Descriptive Statistics for NCE Scores for the SAT 10 Problem Solving Subtest for All Testing Occasions by School

	2004			2005			2006		
School	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
A	27	30.73	16.63	27	39.29	21.87	27	43.89	18.02
B	16	41.18	14.92	16	48.08	26.65	16	48.37	20.74
C	40	41.12	17.65	40	46.79	18.15	40	42.48	18.54

As indicated in Table 23, there was a sizeable difference in the total mean scores for the SAT 10 Problem Solving subtest between 2004 (M = 37.75) and 2005 (M =

44.60). However, in 2006, there was a slight decrease in the total mean (*M* = 44.07). In 2005, the highest scores for Schools B and C were obtained. Other independent variables (gender, socioeconomic status, English proficiency, and special education) were not included in the model because the small sizes of cells caused insufficient power. The small sample size resulted in exclusion of these variables from the multivariate model, but they were examined in the univariate model.

Statistically Significant Findings for Null Hypothesis 2

The null hypothesis of no significant interaction was rejected for the independent variable, school ($F = 3.35$, $df = 4$, $p = .011$), as shown in Table 24. In addition, there was significant interaction between scores and time (testing occasions; $F = 11.18$, $df = 2$, $p < .001$). The data for scores across testing occasions are quadratic in nature and account for 5.2% of the total variability in scores. A quadratic effect is observed when there are more than two time points and is more complex than a linear relationship is found to be. The most complex effect is utilized for interpretation; therefore, the quadratic effect was interpreted. On the other hand, interaction between time (testing occasions) and school is linear, with two points. Descriptive statistics for all independent variables for all testing occasions are presented in Table 25.

Table 24

Multivariate Tests for NCE Scores for the SAT 10 Problem Solving Subtest

Effect	Source	*F*	*df*	Significance	Partial η^2
Time	Pillai's Trace	11.18	2	<.001	.221
Time * School	Pillai's Trace	3.35	4	.011	.077
Time	Quadratic	4.39	1	.039	.052
Time * School	Linear	6.21	2	.003	.134

Table 25

Descriptive Statistics for NCE Scores for the SAT 10 Problem Solving Subtest for Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency and Level of Special Education Needed for All Testing Occasions

	2004			2005			2006		
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Ethnicity									
White	15	40.25	19.57	15	44.21	26.21	15	40.47	20.92
Black	38	36.64	15.76	38	42.68	21.46	38	44.39	18.68
Hispanic	30	37.92	18.53	30	47.22	18.64	30	45.47	18.01
Gender									
Male	40	39.87	16.85	40	46.81	20.98	40	45.24	19.34
Female	43	35.78	17.75	43	42.54	21.57	43	42.99	18.27
Socioeconomic status									
Paid lunch	78	37.22	16.96	78	44.01	21.07	78	43.75	18.72
Free and reduced lunch	5	46.06	23.10	5	53.72	24.75	5	49.14	19.94
Level of English proficiency									
Non-limited English proficiency	55	38.18	16.81	55	43.64	22.40	55	43.36	19.31
Limited English proficiency	28	36.90	18.62	28	46.49	19.08	28	45.46	17.73
Level of special education needed									
Non-special education	74	40.05	16.77	74	48.08	19.69	74	46.29	18.39
Special education	9	18.80	7.93	9	16.01	8.23	9	25.88	9.08

To illustrate the findings, a line graph for mean NCE scores for the SAT 10 Problem Solving subtest across testing occasions was plotted. Figure 3 clearly shows there was an increase in mean NCE scores for the Problem Solving subtest between 2004 and 2005; this result was not surprising because teachers were excited about implementing the MMI. However, between 2005 and 2006, there was no significant change, with the mean score constant at 43. One plausible explanation is that the impact of innovation was greatest at the initial point of exposure. Another explanation would be related to the change in number of mathematics coaches serving MMI-trained teachers. In 2004-2005, there were four mathematics coaches who served 57 MMI-trained kindergarten and third-grade teachers in the eight MMI schools at the time of the study. Total time spent by mathematics coaches in each school averaged 8 hr per week; which was considered to be an ideal scenario in the MMI program. Of the eight MMI schools, three were identified for use in the study based on the criteria stated in chapter 1. In 2005-2006, because of funding issues, the number of mathematics coaches was decreased to two coaches serving 109 MMI-trained teachers. During the summer of 2005, first- and fourth-grade teachers received training in the MMI; therefore, the number of teachers who received mathematics coaching services for the MMI increased from 57 to 109 at the same time that the number of mathematics coaches declined from four to two. Because of the reduction in the coach-to-teacher ratio, the total time spent by mathematics coaches in each school dropped to 8 hr every other week, with less time spent per teacher.

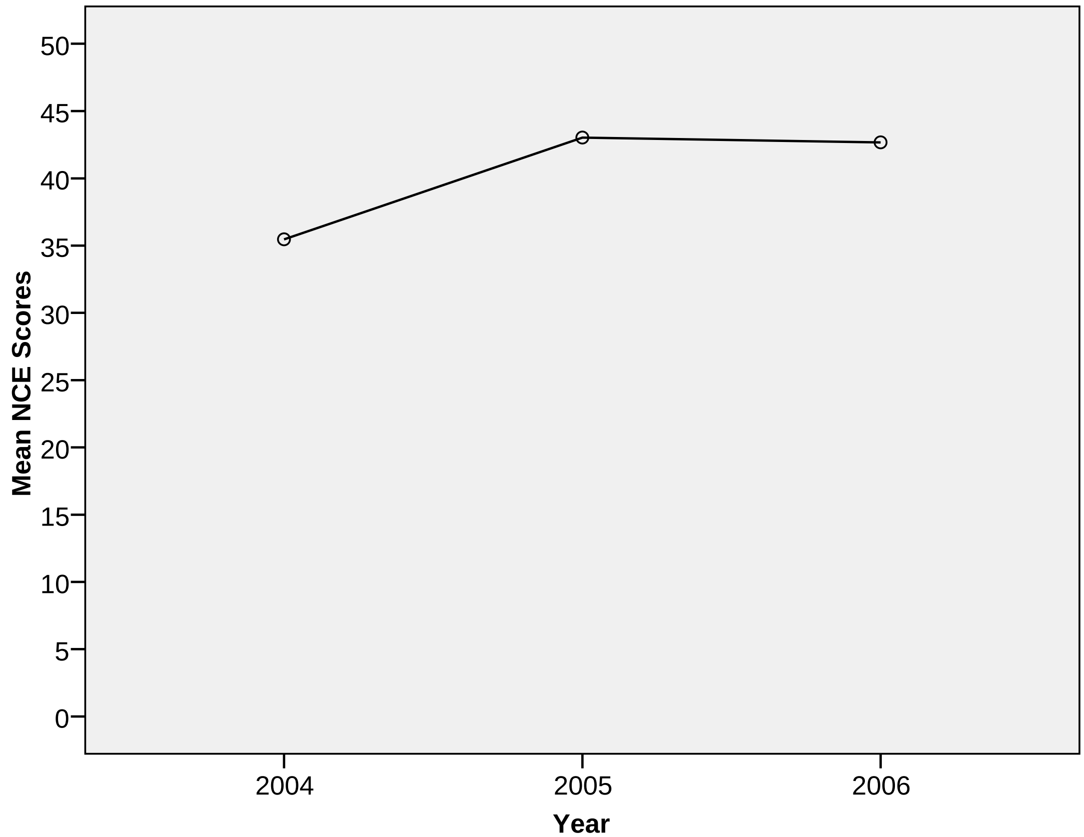

Figure 3. NCE scores for the SAT 10 Problem Solving subtest across testing occasions.

Although the findings showed statistical significance between time (testing occasions) and school and between scores and time (testing occasions) and resulted in rejecting Null Hypothesis 2 for the SAT 10 Problem Solving subtest by school and over time, there was insufficient evidence to reject Null Hypothesis 2 for the other independent variables. This finding can be explained, in part, by the small sample sizes in many of the cells. The sample sizes were too small to detect a statistically significant interaction even if one existed. The variables were examined as main effects in the repeated measures model by using independent *t* tests, paired *t* tests, and one-way ANOVA. Determining whether the differences in means between categories of independent variables were statistically significant was accomplished by conducting independent *t* tests for gender, socioeconomic status, level of English proficiency, and level of special education needed. One-way ANOVA was carried out for ethnicity. Paired *t* tests were conducted to determine whether the differences over time of all categories of independent variables were significant. Results of independent *t* tests are presented in Table 26, those for one-way ANOVA are given in Table 27, and those for paired *t* tests are displayed in Table 28 and 29.

Table 26

Independent t Tests for NCE Scores for the SAT 10 Problem Solving Subtest for Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	Year	*t*	Significance (2-tailed)
Gender	2004	1.073	.286
	2005	0.912	.365
	2006	0.545	.587
Socioeconomic	2004	-1.107	.271
Status	2005	-0.989	.325
	2006	-0.622	.535
Level of English proficiency	2004	0.318	.751
	2005	-0.577	.566
	2006	-0.481	.632
Level of special education needed	2004	3.737	<.001
	2005	4.815	<.001
	2006	3.267	.002

Table 27

One-Way ANOVA for Ethnicity for NCE Scores for the SAT 10 Problem Solving Subtest

Year	*df*	*F*	Significance
2004	2, 80	.23	.79
2005	2, 80	.37	.68
2006	2, 80	.36	.69

Table 28

Paired t Tests for NCE Scores for the SAT 10 Problem Solving Subtest Between 2004 and 2006 for School, Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Level of categorical pairing for 2004-2006	*t*	*df*	Significance
School			
School A	-4.454	26	<.001
School B	-2.133	15	.050
School C	-0.714	39	.480
Ethnicity			
White	-0.062	14	.952
Black	-3.273	37	.002
Hispanic	-2.989	29	.006
Gender			
Male	-2.103	39	.042
Female	-3.786	42	<.001
Socioeconomic status			
Free and reduced lunch	-3.936	77	<.001
Paid lunch	-0.946	4	.398
Level of English proficiency			
Limited English proficiency	-2.599	54	.012
Non-limited English proficiency	-3.403	27	.002
Level of special education needed			
Special education	-2.209	8	.058
Non-special education	-3.614	73	.001

Table 29

Paired t Tests for NCE Scores for the SAT 10 Problem Solving Subtest Between 2004 and 2005 and Between 2005 and 2006 for School

	2004-2005			2005-2006		
School	*t*	*df*	Significance	*t*	*df*	Significance
A	-2.163	26	.040	-1.670	26	.107
B	-1.629	15	.124	-0.069	15	.946
C	-2.815	39	.008	2.028	39	.049

A line graph for Problem Solving NCE scores across schools and across testing occasions was also plotted. As shown in Figure 4, Schools A and B experienced statistically significant gains in NCE scores for the SAT 10 Problem Solving subtest between

2004 and 2006; furthermore, Schools A and C experienced statistically significant gains in the first year of the study, 2004-2005. Last, School C experienced a statistically significant loss in NCE scores for the SAT 10 Problem Solving subtest between 2005 and 2006.

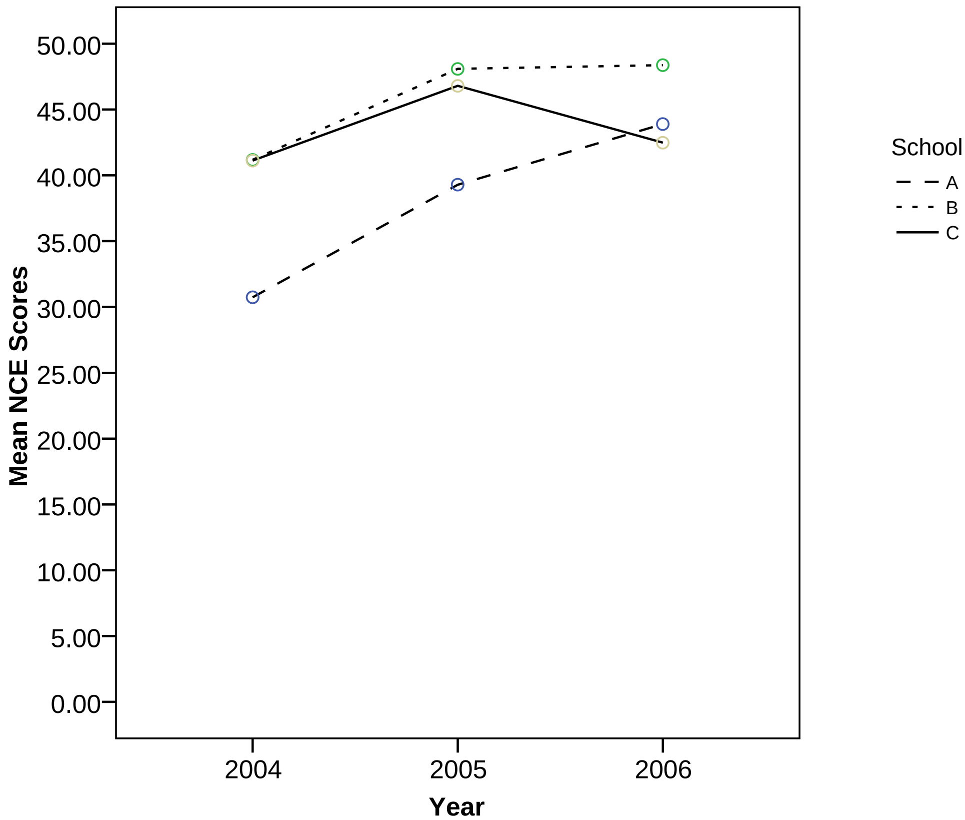

Figure 4. NCE scores for the SAT 10 Problem Solving subtest reported by school.

The following findings are reported:

1. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for Black and Hispanic students were lower than mean NCE scores for White students (M = 36.64 for Black students; M = 37.92 for Hispanic students; M = 40.25 for White students); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for Black

and Hispanic students were higher than mean NCE scores for White students (*M* = 44.39 for Black students; *M* = 45.47 for Hispanic students; *M* = 40.47 for White students). Although the difference in mean NCE scores for the SAT 10 Problem Solving subtest for ethnicity was not found to be statistically significant at baseline ($F = 0.23$, $df = 2, 80$, $p = .79$), or after 2 years of intervention ($F = 0.36$, $df = 2, 80$, $p = .69$), the researcher believes the difference is noteworthy in that mean NCE scores for Black and Hispanic students were lower than mean NCE scores for White students at baseline but were higher after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for Black students over the 3 years of the study showed a statistically significant gain ($t = -3.273$, $df = 37$, $p = .002$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for Hispanic students over the 3 years of the study showed a statistically significant gain ($t = -2.989$, $df = 29$, $p = .006$). Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for White students over the 3 years of the study failed to show a significant gain ($t = -0.062$, $df = 14$, $p = .952$).

2. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for male students were higher than mean NCE scores for female students (*M* = 39.87 for male students; *M* = 35.78 for female students); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for male students remained higher than mean NCE scores for female students (*M* = 45.24 for male students; *M* = 42.99 for female students). The difference in mean NCE scores for the SAT 10 Problem Solving

subtest for gender was not found to be statistically significant at baseline (*t* = 1.073, *df* = 81, *p* = .286), or after 2 years of intervention (*t* = 0.545, *df* = 81, *p* = .587). Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for male students over the 3 years of the study showed a statistically significant gain (*t* = -2.103, *df* = 39, *p* = .042). Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for female students over the 3 years of the study showed a statistically significant gain (*t* = -3.786, *df* = 42, *p* < .001).

3. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for students who received subsidies for lunches were lower than mean NCE scores for students who paid for their lunches (*M* = 37.22 for students who received subsidies for lunches; *M* = 46.06 for students who paid for their lunches); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for students who received subsidies for their lunches remained lower than mean NCE scores for students who paid for their lunches (*M* = 43.75 for students who received subsidies for their lunches; *M* = 49.14 for students who paid for their lunches). The difference in mean NCE SAT 10 Problem Solving subtest scores for socioeconomic status was not found to be statistically significant at baseline (*t* = -1.107, *df* = 81, *p* = .271), or after 2 years of intervention (*t* = -0.622, *df* = 81, *p* = .535). Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for students who received subsidies for their lunches showed a statistically significant gain (*t* = -3.936, *df* = 77, *p* < .001) over the 3 years of the study. Findings of paired *t* tests for mean NCE scores for the SAT 10 Prob-

lem Solving subtest for students who paid for their lunches showed a statistically significant gain (*t* = -0.946, *df* = 4, *p* = .398).

4. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for LEP students were lower than mean NCE scores for non-LEP students (*M* = 36.90 for LEP students; *M* = 38.17 for non-LEP students); after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for LEP students were higher than mean NCE scores for non-LEP students (*M* = 45.46 for LEP students; *M* = 43.36 for non-LEP students). Although the difference in mean NCE scores for the SAT 10 Problem Solving subtest for English proficiency (LEP and non-LEP) was not found to be statistically significant at baseline (*t* = 0.318, *df* = 81, *p* = .751), or after 2 years of intervention (*t* = -0.481, *df* = 81, *p* = .632), the researcher believes it is noteworthy that scores of LEP students started lower than but ended higher than scores of non-LEP students on this subtest after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for LEP students over the 3 years of the study showed a statistically significant gain (*t* = -2.599, *df* = 54, *p* = .012). Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for non-LEP students over the 3 years of the study showed a statistically significant gain (*t* = -3.403, *df* = 27, *p* = .002).
5. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for special education students were significantly lower than mean NCE scores for non-special education students (*M* = 18.80 for special education students; *M* = 40.05 for non-special education students; *t* = 3.737, *df* = 81, *p* < .001); after 2

years of intervention by an MMI-trained teacher, mean NCE scores for Problem Solving for special education students continued to be significantly lower than mean NCE scores for non-special education students (*M* = 25.88 for special education students; *M* = 46.29 for non-special education students; *t* = 3.267, *df* = 81, *p* = .002). Although the significant difference in mean NCE scores for the SAT 10 Problem Solving subtest for special education was expected at both baseline and after 2 years of intervention by an MMI-trained teacher, the researcher believes it is noteworthy that special education students made substantial gains on this subtest after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for special education students over the 3 years of the study failed to show a statistically significant gain (*t* = -2.209, *df* = 8, *p* = .058); however, a trend was noted toward statistical significance. (A trend toward statistical significance is noted when $p \leq 1.0$; G. Abbott, personal communication, May 11, 2007). Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for non-special education students over the 3 years of the study showed a statistically significant gain (*t* = -3.614, *df* = 73, *p* = .001).

Conclusions for Null Hypothesis 2

For Null Hypothesis 2, the null hypothesis of no significant interaction was rejected for the independent variable, school (*F* = 3.35, *df* = 4, *p* = .011). In addition, there was significant interaction between scores and time (testing occasions; *F* = 11.18, *df* = 2, $p < .001$). There was insufficient evidence to reject Null Hypothesis 2 for the other independent variables. This lack of evidence can be explained in part by the small sample size of the cells. Independent *t* tests were significant for special education only ($p < .001$). Paired *t* tests comparing mean NCE scores for the SAT 10 Problem Solving subtest for each category of independent variable at baseline and final testing occasions (2004 and 2006, respectively) were significant for Black, Hispanic, male, female, subsidized, LEP, non-LEP, and non-special education students and were not significant for White, unsubsidized, and special education students.

Findings of Statistical Analysis for Null Hypothesis 3

ANOVA with repeated measures and descriptive statistics were used to examine Null Hypothesis 3: There will be no significant interaction in mathematics achievement as measured by scale scores on the ARMT Math subtest over a 2-year period for students in three high poverty schools from a single school system whose teachers participated in the MMI. As in the first null hypothesis, the independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

On average, there was a slight increase in performance in the ARMT Math subtest scores in 2006 as compared to those from 2005. In both years, School B performed better than Schools A and C did. Table 30 displays descriptive statistics for the ARMT Math scale scores for 2005 and 2006.

Table 30

Descriptive Statistics for Scale Scores for the ARMT Math Subtest

	2005			2006		
School	*n*	*M*	*SD*	*n*	*M*	*SD*
A	29	613.10	40.35	29	625.41	37.56
B	17	629.65	47.17	17	622.59	35.45
C	39	625.87	37.87	39	619.85	36.07

Statistically Significant Findings for Null Hypothesis 3

The null hypothesis of no significant interaction was rejected for the independent variable, school (F = 3.43, df = 2, p = .037), as shown in Table 31. The interaction between testing occasions and school is linear and accounts for approximately 7.7% of variability in the dependent variable. Descriptive statistics for all independent variables for all testing occasions are presented in Table 32.

Table 31

Multivariate Tests for Scale Scores for the ARMT Math Subtest

Effect	Source	*df*	*F*	Significance	Partial η^2
Time	Pillai's Trace	1	0.128	.722	.002
Time * School	Pillai's Trace	2	3.430	.037	.077
Time	Linear	1	0.128	.722	.002
Time * School	Linear	2	3.430	.037	.077

Table 32

Descriptive Statistics for Scale Scores for the ARMT Math Subtest for Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed for Both Testing Occasions

	2005			2006		
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*
Ethnicity						
White	14	617.64	40.50	14	612.43	34.80
Black	40	622.10	43.00	40	621.85	37.34
Hispanic	31	624.58	39.00	31	627.32	35.26
Gender						
Male	41	624.83	39.61	41	625.34	35.14
Female	44	619.89	42.11	44	619.45	37.18
Socioeconomic status						
Paid lunch	79	621.29	40.71	79	621.05	34.64
Free and reduced lunch	6	635.17	42.95	6	638.67	37.00
Level of English proficiency						
Non-limited English proficiency	56	621.14	41.41	56	619.27	36.18
Limited English proficiency	29	624.45	41.00	29	628.14	35.90
Level of special education needed						
Non-special education	76	628.03	38.23	76	626.32	35.63
Special education	9	573.67	27.63	9	588.33	18.53

To further explain the trend of performance in procedures between testing occasions, line graphs were plotted. As shown in Figure 5, there was a constant and linear increase in means for the ARMT Math scores between 2005 and 2006.

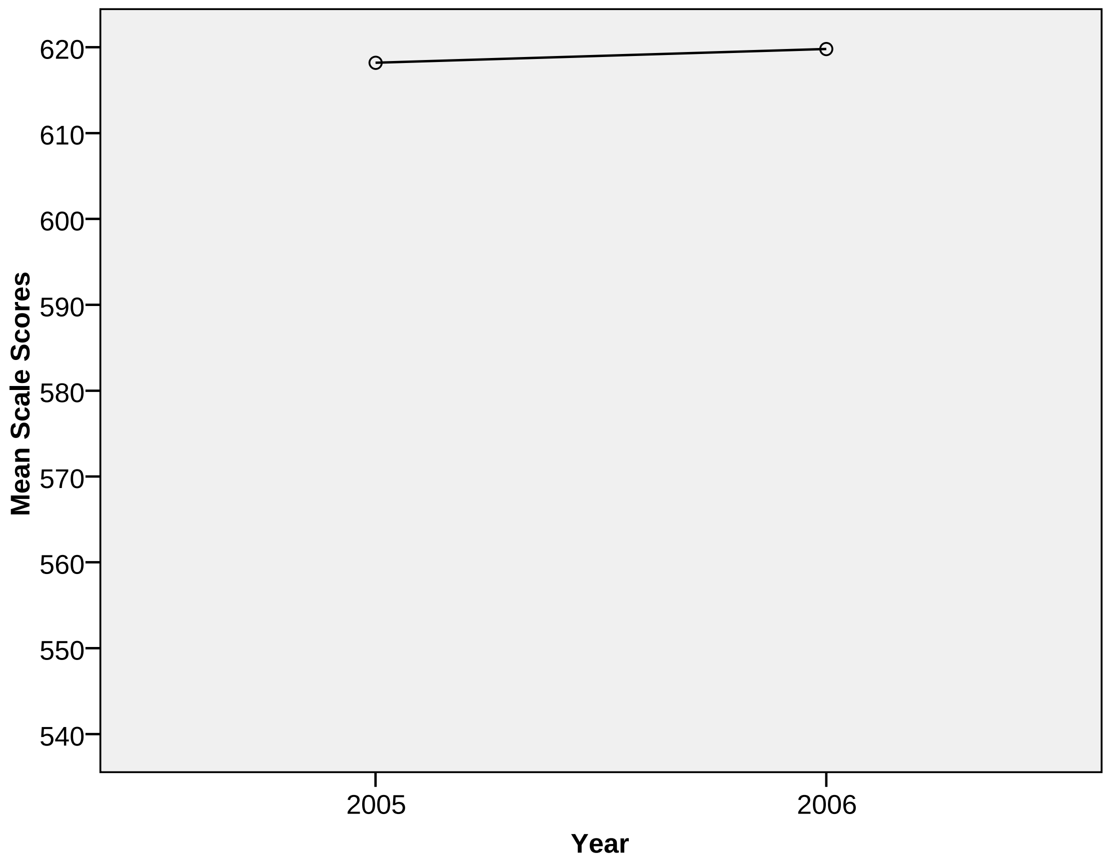

Figure 5. Scale scores for the ARMT Math subtest between testing occasions.

Although the findings for the ARMT Math subtest scores showed a statistically significant interaction effect between time (testing occasions) and school, there was insufficient evidence to reject Null Hypothesis 3 for the other independent variables (gender, socioeconomic status, English proficiency, and special education). This finding can be explained, in part, by the small sample sizes in many of the cells. The sample sizes were too small to detect a statistically significant interaction even if one existed. Thus, the variables were examined as main effects in the repeated measures model. Independent *t* tests, paired *t* tests, and one-way ANOVA were used to explore the data. To determine whether the differences in means between categories of independent variables were statistically significant, independent *t* tests were conducted for gender, socioeconomic status, level of English proficiency, and level of special education needed. One-way ANOVA was carried out for ethnicity. To determine whether the differences over time of all categories of independent variables were significant, paired *t* tests were conducted. Results of independent *t* tests are presented in Table 33, those for one-way ANOVA are given in Table 34, and those for paired *t* tests are displayed in Table 35.

Table 33

Independent t tests for Scale Scores for the ARMT Math Subtest for Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	Year	*t*	Significance (2-tailed)
Gender	2005	0.556	.579
	2006	0.749	.456
Socioeconomic status	2005	-0.802	.425
	2006	-1.154	.252
Level of English proficiency	2005	-0.353	.725
	2006	-1.075	.286
Level of special education needed	2005	4.129	<.001
	2006	3.133	.002

Table 34

One-way ANOVA for Ethnicity for Scale Scores for the ARMT Math Subtest

Year	*df*	*F*	Significance
2005	2, 82	0.138	.872
2006	2, 82	0.822	.443

Table 35

Paired t Tests for Scale Scores for the ARMT Math Subtest Between 2005 and 2006 for School, Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Level of categorical pairing for 2004-2006	*t*	*df*	Significance
School			
School A	-2.948	28	.006
School B	0.994	16	.335
School C	1.215	38	.232
Ethnicity			
White	1.036	13	.319
Black	0.048	39	.962
Hispanic	-0.539	30	.594
Gender			
Male	-0.130	40	.897
Female	0.088	43	.930
Socioeconomic status			
Free and reduced lunch	0.074	78	.941
Paid lunch	-0.242	5	.818
Level of English proficiency			
Limited English proficiency	-0.684	28	.499
Non-limited English proficiency	0.481	55	.632
Level of special education needed			
Special education	-2.334	8	.048
Non-special education	0.503	75	.617

A line graph was used to display school performance comparisons between the scale scores for the ARMT Math subtest in 2005 and 2006. As shown in Figure 6, School A experienced a statistically significant gain in scale scores for the ARMT Math subtest during the 2nd of the study (2005-2006). Schools B and C experienced a decline in test scores for the ARMT Math subtest that was not found to be statistically significant. The researcher expected this finding because of the very high intercorrelation (.99) between the ARMT Math subtest and the SAT 10 Problem Solving subtest that was reported in the *Alabama Reading and Math Test Technical Manual* (2005).

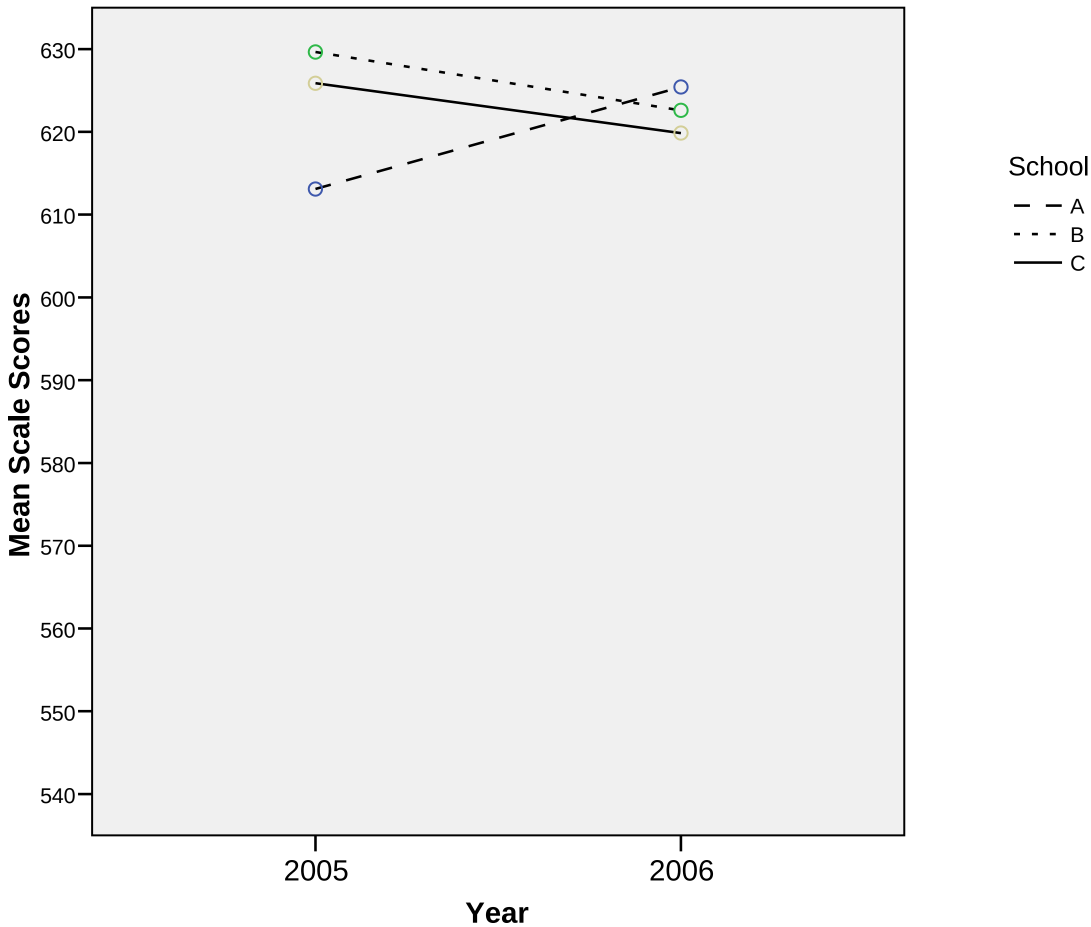

Figure 6. Scale score for the ARMT Math subtest reported by school.

The following findings are reported:

1. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for Black and Hispanic students were higher than mean scale scores for White students (*M* = 622.10 for Black students; *M* = 624.58 for Hispanic students; *M* = 617.64 for White students); in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for Black and Hispanic students continued to be higher than mean scale scores for White students (*M* = 621.85 for

Black students; *M* = 627.32 for Hispanic students; *M* = 612.43 for White students). The difference in mean scale scores for the ARMT Math subtest for ethnicity was not found to be statistically significant after the first year of intervention (*F* = 0.138, *df* = 2, 82, *p* = .872), or after the 2nd year of intervention (*F* = 0.822, *df* = 2, 82, *p* = .443). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for Black students between 2005 and 2006 did not show a statistically significant gain (*t* = 0.048, *df* = 39, *p* = .962). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for Hispanic students between 2005 and 2006 did not show a statistically significant gain (*t* = -0.539, *df* = 30, *p* = .594). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for White students over the 3 years of the study failed to show a significant gain (*t* = 1.036, *df* = 13, *p* = .319).

2. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for male students were higher than mean scale scores for female students (*M* = 624.83 for male students; *M* = 619.89 for female students); in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for male students continued to be higher than mean scale scores for female students (*M* = 625.34 for male students; *M* = 619.45 for female students). The difference in mean scale scores for the ARMT Math subtest for gender was not found to be statistically significant after the first year of intervention (*t* = 0.556, *df* = 83, *p* = .579), or after the 2nd year of intervention (*t* = 0.749, *df* = 83, *p* = .456). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for male students between 2005 and 2006 failed to show a statistically

significant gain (t = -0.130, df = 40, p = .897). Findings of paired t tests for mean scale scores for the ARMT Math subtest for female students between 2005 and 2006 failed to show a statistically significant gain (t = 0.088, df = 43, p = .930).

3. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for students who received subsidies for lunches were lower than scores for students who paid for their lunches (M = 621.29 for students who received subsidies for lunches; M = 635.17 for students who paid for their lunches); in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for students who received subsidies for their lunches continued to be lower than mean scale scores for students who paid for their lunches (M = 621.05 for students who received subsidies for their lunches; M = 638.67 for students who paid for their lunches). The difference in mean scale scores for the ARMT Math subtest for socioeconomic status was not found to be statistically significant after the first year of intervention (t = -0.802, df = 83, p = .425), or after the 2nd year of intervention (t = -1.154, df = 83, p = .242). Findings of paired t tests for mean scale scores for the ARMT Math subtest for students who received subsidies for their lunches failed to show a statistically significant gain (t = 0.074, df = 78, p = .941) between 2005 and 2006. Findings of paired t tests for mean scale scores for the ARMT Math subtest for students who paid for their lunches failed to show a statistically significant gain (t = -0.242, df = 5, p = .818) between 2005 and 2006.

4. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for LEP students were higher than mean scale scores for non-LEP students (*M* = 624.45 for LEP students; *M* = 621.14 for non-LEP students); in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for LEP students continued to be higher than mean scale scores for non-LEP students (*M* = 628.14 for LEP students; *M* = 619.27 for non-LEP students). The difference in mean scale scores for the ARMT Math subtest for English proficiency was not found to be statistically significant after the first year of intervention ($t = -0.353$, $df = 83$, $p = .725$), or after the 2nd year of intervention ($t = -1.075$, $df = 83$, $p = .286$). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for LEP students between 2005 and 2006 failed to show a statistically significant gain ($t = -0.684$, $df = 28$, $p = .499$). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for non-LEP students between 2005 and 2006 failed to show a statistically significant gain ($t = 0.481$, $df = 55$, $p = .632$).
5. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for special education students were significantly lower than mean scale scores for non-special education students (*M* = 573.67 for special education students; *M* = 628.03 for non-special education students; $t = 4.129$, $df = 83$, $p < .001$); after 2 years of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for special education students continued to be significantly lower than mean scale scores for non-special education students (*M* = 588.33 for special education

students; *M* = 626.32 for non-special education students; *t* = 3.133, *df* = 83, *p* = .002). Although the significant difference in mean scale scores for the ARMT Math subtest for special education was expected at both the first and 2nd years of intervention by an MMI-trained teacher, the researcher believes it is noteworthy that special education students made substantial gains on this subtest after 2 years of intervention. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for special education students between 2005 and 2006 showed a statistically significant gain (*t* = -2.334, *df* = 8, *p* = .048). Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for non-special education students between 2005 and 2006 failed to show a statistically significant gain (*t* = -0.503, *df* = 75, *p* = .617).

Conclusions for Null Hypothesis 3

Findings for Null Hypothesis 3 showed statistical significance between testing occasions for the ARMT Math subtest across all schools (*F* = 3.430, *df* = 2, *p* = .037), but there was insufficient evidence to reject Null Hypothesis 3 for the other independent variables. Independent *t* tests were significant for special education at both testing occasions. Paired *t* tests comparing mean scale scores for the ARMT Math subtest for each category of independent variable between testing occasions in 2005 and 2006 were significant for special education students, only.

Summary

In analyzing the study data for each of the null hypotheses, the researcher noted a statistical significance for the interaction between time (testing occasions) and schools.

Additionally, there was a significant interaction between mean NCE scores for the SAT 10 Procedures and Problem Solving subtests and time (testing occasions). Important results from the study are that two findings were consistently identified:

1. Findings for Null Hypotheses 2 and 3 showed that LEP students consistently outperformed non-LEP students on the ARMT Math subtest and the SAT 10 Problem Solving subtest after 2 years of intervention by MMI-trained teachers.
2. Findings for Null Hypotheses 1, 2, and 3 showed that Black and Hispanic students outperformed White students on the ARMT Math subtest and the SAT 10 Procedures and Problem Solving subtests after 2 years of intervention by MMI-trained teachers.

CHAPTER 5

SUMMARY, DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

This chapter provides a brief summary of the entire study, including the primary objectives, significance of the study, theoretical framework, study methodology, and major findings. Implications of the findings for the profession and recommendations for practice and future research are also presented.

Summary

A deep understanding of mathematics with well-developed quantitative literacy skills is critical for 21st-century students. Given the relatively low international rankings of U.S. students in mathematics despite the development of NCTM standards and numerous attempts at reform, the ways in which students can most effectively learn to be mathematical thinkers and quantitatively literate citizens are not clear. It has been suggested that immediate reform efforts offer teachers the opportunity to transform their learning about mathematics and will lead to effective teaching practices that prepare students to be quantitatively literate citizens of the world.

The U.S. educational system has repeatedly implemented reform efforts in the teaching of mathematics, with very little change in classroom performance (Leinwand, 2000). According to analysis and conclusions from the TIMSS video study (Stigler & Hiebert, 1997), U.S. teachers often use traditional teaching methods to show students what to do without developing specific mathematics concepts to foster a deeper level of understanding. Problems assigned for practice are similar to the worked examples (Stigler & Hiebert, 1999). The problem that faces the United States is as follows: How can U.S. preservice and in-service teachers make dramatic and fundamental changes in classroom practice that lead to students' acquiring a deeper understanding of mathematics and improving their achievement in mathematics?

Improving mathematical achievement involves changing both the instructional practices of teachers and the learning environment for students. According to the NCTM Standards (2000), "mathematical expertise is found not only in factual knowledge, but also in the strategic decisions the student makes while solving problems and in their ability to communicate ideas about the problems with others" (Hiebert, 2003, p. 13).

The purpose of the study was to determine the effectiveness of one professional development intervention, the MMI, as implemented in three high poverty schools in a single school district. Changes in student achievement over a 3-year period were examined by using NCE scores from the SAT 10 Procedures and Problem Solving subtests and scale scores from the ARMT Math subtest. The independent variables in the study were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended. The dependent variables were NCE scores for the SAT 10 Procedures and Problem Solving subtests and scale scores for the ARMT Math subtest.

There were three null hypotheses in the study:

1. There will be no significant interaction in mathematics achievement as measured by NCE scores on the Procedures subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The independent variables examined

were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

2. There will be no significant interaction in mathematics achievement as measured by NCE scores on the Problem Solving subtest of the SAT 10 over a 3-year period for students in three high poverty schools in a single school system whose teachers had participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

3. There will be no significant interaction in mathematics achievement as measured by scale scores on the ARMT Math subtest over a 2-year period for students in three high poverty schools from a single school system whose teachers participated in the MMI. The independent variables examined were (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended.

Although many professional development programs have been devised in an attempt to improve mathematics skills of high poverty students, results from those studies have yielded equivocal findings regarding success of these reform initiatives. Because the MMI has been demonstrated to be successful in improving student achievement of this targeted group, this study has tested the effectiveness of the MMI in an additional location with a similar group of students (Van Hanegan et al., 2004).

Findings from this study are useful to teachers, principals, curriculum specialists, mathematics specialists, professional development coordinators, and superintendents in the selection and implementation of nontraditional programs for teaching mathematics. Because the MMI provides instruction that improved learning outcomes for the students in the present study, other school systems may also benefit from using this evidence-based method for teaching mathematics in multiple settings. Efforts such as using the MMI may fill the need for individuals in the 21st-century learning environment who are quantitatively literate, both as competent citizens and as competent employees.

The review of literature focused on the continuing need for mathematics reform and on the steps that are being taken in the United States, as well as in the state of Alabama, to improve student achievement in mathematics. Literature related to four broad topics relevant to the study was reviewed. These topics were (a) mathematics reform efforts, which include national and international studies; (b) U.S. efforts to improve mathematics achievement; (c) exemplary and promising programs; and (d) an overview of the MMI.

This study included a description of the research design, intervention, sample, and data employed in the study. In addition, descriptions of the criteria used for selecting the sample and of the population demographics from which the samples were derived were provided. From the population of second-, third-, and fourth- grade students of the study schools, 89 students with scores for each of the three testing occasions (2004, 2005, and 2006) were enrolled in the study. NCE scores from the SAT 10 Procedures and Problem Solving subtests from 2004 to 2006 and scale scores from the ARMT Math subtest from 2005 to 2006 were gathered. The effects on student achievement that result when the student is taught by an MMI-trained teacher were explored by examining the relationship of mathematics achievement scores to the independent variables of (a) ethnic background, (b) gender, (c) socioeconomic status, (d) level of English proficiency, (e) level of special education needed, and (f) school student attended. The three null hypotheses were examined with the use a factorial ANOVA with repeated measures.

MMI Professional Development

The ensuing description outlines the professional development provided to the teachers of student participants in the present study. The MMI training consists of the following elements: Consultants who have been trained in MMI strategies and who are employed by Mathworks under the direction of Dr. Honi Bamberger conduct the professional development workshops. Mathematics coaches and experienced MMI teachers co-present with the Mathworks presenters. Included in the workshops are opportunities for participants to individually practice their newly learned skills by teaching students enrolled in summer school. Feedback is provided to each teacher by a peer observer for five mornings during the midpoint of this 2-week intervention. Each day for 5 days, the observer facilitates a reflective conversation with the teacher. In addition, principals and teachers have faculty planning time during the workshop to restructure their curriculum to include the MMI strategies.

Four unique strategies that are daily routines for the MMI teachers and their students are as follows: Daily Data activities, Early Bird math activities, and THINK-PAIR-SHARE and other management activities. Daily Data activities that include Venn diagrams and graphs are used to engage students in activities that make connections between real life and mathematics. Early Bird math activities are short, 10 min, independent, hands-on activities that reinforce mathematics skills and concepts. THINK-PAIR-

SHARE, a cooperative learning strategy, is a way to manage discussions with students and provide them with thinking time. "Equity sticks" are an additional strategy for managing any discussion with students.

Specific strategies expected of teachers in the MMI workshops are congruent with best practices in mathematics (NCTM, 2000) and include the following strategies:

1. Plan instruction that is based on students' prior knowledge.
2. Facilitate additional learning opportunities by allowing adequate time for students to think and listen to each other.
3. Use questioning strategies that promote deep levels of understanding by the students. Teachers ask questions such as "How do you know that?" and "Will that work every time?" and "Did anyone get a different answer or work the problem another way?" In addition, students provide answers that justify mathematics concepts.
4. Focus on students' thinking and problem-solving strategies instead of only on obtaining the correct answer.
5. Reinforce skills that enhance estimation and mental math.
6. Reinforce computational skills by connecting problem solving to real-world applications.
7. Integrate mathematics throughout and across the curriculum.
8. Assess student knowledge and spontaneously modify instruction (NCTM, 2000).

In an MMI classroom, desks are grouped together for collaborative problem solving, and mathematic manipulatives are available for easy access. Calendar Math is evident in all classrooms. Student writing samples that show their mathematical thinking are displayed throughout the school building.

After the 2-week training, additional training opportunities are provided, including a 3-day advanced training workshop that teachers attend the summer after their 2-week training. Teachers are paid a stipend to attend workshops and are given several hundred dollars' worth of preselected grade-level-specific materials that are similar to the materials used in the 2-week summer institute to use in their classrooms; these materials include math manipulatives. Teachers attend a priming workshop during the first summer, a 2-week institute during the second summer, and a 3-day advanced institute during the third summer.

Although several unique features of the MMI are described in the literature, a specific feature of interest is that of the inclusion of mathematics coaches hired to assist teachers in the classroom. Mathematics coaches focus on teachers who have been through the 2-week professional training. The coaches provide coaching and mentoring support to teachers, assist in planning, demonstrate lessons, co-plan, co-teach, conduct observations, conduct monthly mathematics presentation to faculties, coordinate MMI plans with principals, coordinate follow-up classroom visits with consultants, participate in additional training, and keep records. Teachers from all eight of the MMI schools meet together monthly and are provided with release time for planning sessions conducted within their schools. Additionally, the MMI consultants provide quarterly visits to the classrooms of the MMI teacher participants to provide feedback.

To support, enhance, and sustain the MMI innovation, teachers may apply to be selected to participate in the rigorous Teacher Leadership Program. The 3-year professional development program focuses on mathematics pedagogy, content, and leadership.

In the 2nd and 3rd years of the program, participating teachers immerse themselves in a school-based leadership practicum.

To assist with implementing and sustaining the MMI, additional structures are necessary. One structure for both implementing the MMI with integrity and sustaining the MMI is the schools' Implementation Committee. High implementers of the MMI are grade-level teachers who have integrated the MMI strategies into most of their teaching. These teachers are selected by the principals in the winter after their initial 10-day professional development. The primary functions of committee members are (a) to serve as a sounding board for coaches and district liaisons, (b) to be highly informed about the MMI, (c) to be advocates of the MMI, (d) to administer quarterly problems to students and to participate in discussion about student thinking related to these quarterly problems, and (e) to participate in professional learning teams (S. Pruet, personal communication, August 13, 2006).

The population for this study consisted of 89 students who were enrolled in three MMI schools in the selected school district; were second-grade students during the 2003-2004 school year; and remained in the same school through the 2005-2006 school year, when they were fourth-grade students. Enrollment in these three MMI schools includes Caucasian, African American, and Hispanic students. Combined demographics for the schools are summarized in Table 36.

Table 36

Combined Demographics for Elementary Schools A, B, and C by Ethnicity, Gender, Socioeconomic Status, Level of English Proficiency, and Level of Special Education Needed

Variable	*f*	% of sample
Ethnicity		
White	16	18.0
Black	42	47.2
Hispanic	31	34.8
Gender		
Male	43	48.3
Female	46	51.7
Socioeconomic status		
Paid lunch	6	6.7
Free and reduced lunch	83	93.3
Level of English proficiency		
Non-limited English proficiency	60	67.4
Limited English proficiency	29	32.6
Level of special education needed		
Non-special education	78	87.6
Special education	11	12.4

These schools were selected because they are composed of high poverty students, with over 80% receiving free or reduced-price lunch. Principals of these schools agreed to administer the SAT 10 to the school's second-grade students; the results of this testing provided pre-MMI baseline data. The demographics included frequency and sample size for gender, ethnicity, level of English proficiency and level of special education needed.

To arrive at findings for the null hypotheses, data analyses using ANOVA with repeated measures were performed. Assumptions of this statistical procedure were discussed in chapter 4, included assumptions of sphericity, an index of variability among schools. Data from the 89 students enrolled in the three MMI schools were analyzed by using the SPSS (Version 15.0).

Findings of the Null Hypotheses

The purpose of chapter 4 was to report findings for the three null hypotheses. ANOVA with repeated measures was used to test the hypotheses. For Null Hypothesis 1, the null hypothesis of no significant interaction was rejected for the independent variable, school. These findings showed statistical significance between time (testing occasions) and school. Schools A and B experienced statistically significant gains in NCE scores for the SAT 10 Procedures subtest between 2004 and 2006; furthermore, School A experienced statistically significant gains at each testing occasion. School C experienced gains in SAT 10 Procedures subtest NCE scores that were not found to be statistically significant. In addition, there was significant interaction between scores and time (testing occasions). There was insufficient evidence to reject Null Hypothesis 1 for the other independent variables. This finding can be explained in part by the small sample sizes in many of the cells.

Findings for each independent variable in Null Hypothesis 1 are as follows:

1. At baseline, mean NCE scores for the SAT 10 Procedures subtest for Black and Hispanic students were similar to mean NCE scores for White students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Procedures subtest for Black and Hispanic students were higher than mean NCE scores for White students. On the basis of results of one-way ANOVA, the difference in mean NCE scores for the SAT 10 Procedures for ethnicity was not found to be statistically significant at baseline, or after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for Black students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean

NCE scores for the SAT 10 Procedures subtest for Hispanic students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for White students over the 3 years of the study failed to show a significant gain.

2. At baseline, mean NCE scores for the SAT 10 Procedures subtest for female students were slightly higher than mean NCE scores for male students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for female students were considerably higher than mean NCE scores for male students. On the basis of results of independent *t* tests, the difference in mean NCE scores for the SAT 10 Procedures subtest for gender was not found to be statistically significant at baseline, or after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for male students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for female students over the 3 years of the study showed a statistically significant gain.
3. At baseline, mean NCE scores for the SAT 10 Procedures subtest for students who received subsidies for lunches were similar to mean NCE scores for students who paid for their lunches; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for Procedures subtest for students who paid for their lunches remained higher than mean NCE scores for students who received subsidies for their lunches. On the basis of independent *t* tests, the difference in mean scores for the SAT 10 Procedures subtest for socioeconomic status was not found to be statistically significant at baseline, or after 2

years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for students who received subsidies for their lunches showed a statistically significant gain over the 3 years of the study. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for students who paid for their lunches showed a statistically significant gain.

4. At baseline, mean NCE scores for the SAT 10 Procedures subtest for LEP students were lower than mean NCE scores for non-LEP students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Procedures subtest for LEP students remained lower than mean NCE scores for non-LEP students. On the basis of results of independent *t* tests, the difference in mean NCE scores for the SAT 10 Procedures subtest for English proficiency was not found to be statistically significant at baseline, or after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for LEP students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for non-LEP students over the 3 years of the study showed a statistically significant gain.
5. At baseline, mean NCE scores for the SAT 10 Procedures subtest for special education students were lower than mean NCE scores for non-special education students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Procedures subtest for special education students continued to be lower than mean NCE scores for non-special education students. On the basis of results of independent *t* tests, the difference in mean

NCE scores for the SAT 10 Procedures subtest for special education was found to be statistically significant at baseline, and after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for special education students over the 3 years of the study failed to showed statistically significant gain; however, a trend was noted toward statistical significance. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for non-special education students over the 3 years of the study showed a statistically significant gain.

For Null Hypothesis 2, the null hypothesis of no significant interaction was rejected for the independent variable, school. Schools A and B experienced statistically significant gains in NCE scores for the SAT 10 Problem Solving subtest between 2004 and 2006; furthermore, Schools A and C experienced statistically significant gains in the first year of the study, 2004-2005. Last, School C experienced a statistically significant loss in NCE scores for the SAT 10 Problem Solving subtest between 2005 and 2006. In addition, there was significant interaction between scores and time (testing occasions). There was insufficient evidence to reject Null Hypothesis 2 for the other independent variables. This lack of evidence can be explained in part by the small sample size of the cells.

Findings for each independent variable are described below:

1. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for Black and Hispanic students were lower than mean NCE scores for White students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for Black and Hispanic students were higher than mean NCE scores for White students. Although the

difference in mean NCE scores for the SAT 10 Problem Solving subtest for ethnicity was not found to be statistically significant at baseline, or after 2 years of intervention, the researcher believes the difference is noteworthy in that mean NCE scores for Black and Hispanic students were lower than mean NCE scores for White students at baseline but were higher after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for Black students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for Hispanic students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Procedures subtest for White students over the 3 years of the study failed to show a significant gain.

2. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for male students were higher than mean NCE scores for female students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for male students remained higher than mean NCE scores for female students. The difference in mean NCE scores for the SAT 10 Problem Solving subtest for gender was not found to be statistically significant at baseline, or after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for male students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for female students over the 3 years of the study showed a statistically significant gain.

3. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for students who received subsidies for lunches were lower than mean NCE scores for students who paid for their lunches; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for students who received subsidies for their lunches remained lower than mean NCE scores for students who paid for their lunches. The difference in mean NCE SAT 10 Problem Solving subtest scores for socioeconomic status was not found to be statistically significant at baseline, or after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for students who received subsidies for their lunches showed a statistically significant gain over the 3 years of the study. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for students who paid for their lunches showed a statistically significant gain.
4. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for LEP students were lower than mean NCE scores for non-LEP students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for the SAT 10 Problem Solving subtest for LEP students were higher than mean NCE scores for non-LEP students. Although the difference in mean NCE scores for the SAT 10 Problem Solving subtest for English proficiency was not found to be statistically significant at baseline, or after 2 years of intervention, the researcher believes it is noteworthy that scores of LEP students started lower than but ended higher than non-LEP students on this subtest after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for

the SAT 10 Problem Solving subtest for LEP students over the 3 years of the study showed a statistically significant gain. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for non-LEP students over the 3 years of the study showed a statistically significant gain.

5. At baseline, mean NCE scores for the SAT 10 Problem Solving subtest for special education students were significantly lower than mean NCE scores for non-special education students; after 2 years of intervention by an MMI-trained teacher, mean NCE scores for Problem Solving for special education students continued to be significantly lower than mean NCE scores for non-special education students. Although the significant difference in mean NCE scores for the SAT 10 Problem Solving subtest for special education was expected at both baseline and after 2 years of intervention by an MMI-trained teacher, the researcher believes it is noteworthy that special education students made substantial gains on this subtest after 2 years of intervention. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for special education students over the 3 years of the study failed to showed a statistically significant gain; however, a trend was noted toward statistical significance. Findings of paired *t* tests for mean NCE scores for the SAT 10 Problem Solving subtest for non-special education students over the 3 years of the study showed a statistically significant gain.

Findings for Null Hypothesis 3 showed statistical significance between testing occasions for the ARMT Math subtest across all schools, but there was insufficient evidence to reject Null Hypothesis 3 for the other independent variables. School A experienced a statistically significant gain in scale scores for the ARMT Math subtest during the second year of the study (2005-2006), and Schools B and C experienced a decline in test scores for the ARMT Math subtest that was not found to be statistically significant. It is not possible to evaluate the overall gain in mathematics learning by using the ARMT Math subtest because this test did not provide baseline data for the study.

Findings for each independent variable are as follows:

1. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for Black and Hispanic students were higher than mean scale scores for White students; in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for Black and Hispanic students continued to be higher than mean scale scores for White students. The difference in mean scale scores for the ARMT Math subtest for ethnicity was not found to be statistically significant after the first year of intervention, or after the 2nd year of intervention. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for Black students between 2005 and 2006 did not show a statistically significant gain. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for Hispanic students between 2005 and 2006 did not show a statistically significant gain. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for White students over the 3 years of the study failed to show a significant gain.
2. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for male students were higher than mean scale scores for female students; in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math sub-

test for male students continued to be higher than mean scale scores for female students. The difference in mean scale scores for the ARMT Math subtest scores for gender was not found to be statistically significant after the first year of intervention, or after the 2nd year of intervention. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for male students between 2005 and 2006 failed to show a statistically significant gain. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for female students between 2005 and 2006 failed to show a statistically significant gain.

3. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for students who received subsidies for lunches were lower than scores for students who paid for their lunches; in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for students who received subsidies for their lunches continued to be lower than mean scale scores for students who paid for their lunches. The difference in mean scale scores for the ARMT Math subtest for socioeconomic status was not found to be statistically significant after the first year of intervention, or after the 2nd year of intervention. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for students who received subsidies for their lunches failed to show a statistically significant gain between 2005 and 2006. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for students who paid for their lunches failed to show a statistically significant gain between 2005 and 2006.

4. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for LEP students were higher than mean scale scores for non-LEP students; in 2006, after the 2nd year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for LEP students continued to be higher than mean scale scores for non-LEP students. The difference in mean scale scores for the ARMT Math subtest for English proficiency was not found to be statistically significant after the first year of intervention, or after the 2nd year of intervention. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for LEP students between 2005 and 2006 failed to show a statistically significant gain. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for non-LEP students between 2005 and 2006 failed to show a statistically significant gain.
5. In 2005, after the first year of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for special education students were significantly lower than mean scale scores for non-special education students; after 2 years of intervention by an MMI-trained teacher, mean scale scores for the ARMT Math subtest for special education students continued to be significantly lower than mean scale scores for non-special education students. Although the significant difference in mean scale scores for the ARMT Math subtest for special education was expected at both the first and 2nd years of intervention by an MMI-trained teacher, the researcher believes it is noteworthy that special education students made substantial gains on this subtest after 2 years of intervention. Findings of paired *t* tests for mean scale scores for

the ARMT Math subtest for special education students between 2005 and 2006 showed a statistically significant gain. Findings of paired *t* tests for mean scale scores for the ARMT Math subtest for non-special education students between 2005 and 2006 failed to show a statistically significant gain.

In each of the null hypotheses, the null hypothesis of no significant interaction was rejected for the independent variable, school. In addition, for Null Hypotheses 1 and 2, there was significant interaction between scores and time (testing occasions). In each case, there was insufficient evidence to reject the null hypothesis for the other independent variables.

Discussion

Results of the present study were congruent with findings of Peske & Haycock (2006), who stated that a high-quality, well-trained teacher is one resource that minority students need in order to reach their potential. Fosnot & Dolk (2002) stated that "children will rise to the challenge when they are able to construct meaning for the world in which they live" (p. 18). One of the key components of the MMI is the daily focus on students' thinking and constructing meaning rather than on their arriving at the correct answer. The MMI-trained teachers are taught the importance of understanding the thinking of their students by using strategies such as asking probing questions that prompt students to explain their thinking; displaying students' work containing narratives by the students that explain their thinking about mathematics procedures; and utilizing ME TOO, POINT OF INTEREST, and THINK-PAIR-SHARE. Teachers learn that all of these strategies are types of formative assessments that help them deeply understand the thinking of their students. The ability to help students articulate their thinking in mathematics class transfers to other areas of study.

The study findings demonstrated that students of MMI-trained teachers improved their NCE scores for the SAT 10 Procedures and Problem Solving subtests from baseline scores (2004) to the end of the study (2006). This study corroborated findings from Bamberger (2005) about MMI training of teachers of minority students. Over the 3 years of the study, mean NCE scores for the SAT 10 Procedures subtest rose significantly for each year of the study; however, mean NCE scores for the SAT 10 Problem Solving subtests rose significantly only in the first year of the study (2004-2005). Scores for the SAT 10 Problem Solving subtest increased in the 2nd year of the study (2005-2006); however, the increase was not found to be statistically significant. This pattern for the SAT 10 Problem Solving subtest scores of significant increase in the first year of the study and a non-statistically significant increase in the 2nd year of the study may well have been seen in the ARMT math subtest if there had been baseline data to use for comparison. The extremely high correlation between the ARMT Math Subtest and the problem solving subtest (.99) provides the basis of this assumption. Mean scale scores for the ARMT Math subtest increased during the 2nd year of the study (2005-2006); like the increase found for the SAT 10 Problem Solving subtest this increase was not found to be statistically significant. While considering potential explanations for this finding, the researcher examined the ratio of mathematics coaches to MMI-trained teachers throughout the study.

In 2004-2005, there were four mathematics coaches who served 57 MMI-trained kindergarten and third-grade teachers in the eight MMI schools at the time of the study. Total time spent by mathematics coaches in each school averaged 8 hr per week, which was considered to be an ideal scenario in the MMI program. Of the eight MMI schools, three were identified for use in the study on the basis of criteria stated in chapter 1. In 2005-2006, funding issues caused the number of mathematics coaches to be decreased to two coaches serving 109 MMI-trained teachers. During the summer of 2005, first- and fourth-grade teachers received training in the MMI; therefore, the number of teachers who received mathematics coaching services for the MMI increased from 57 to 109 at the same time that the number of mathematics coaches declined from four to two. Because of the reduction in the coach-to-teacher ratio, the total time spent by mathematics coaches in each school dropped to 8 hr every other week, with less time spent per teacher.

Because support and accountability for the MMI-trained teachers by the mathematics coaches are integral aspects of the MMI, the shift in the ratio of coaches to teachers may well have influenced the results of the study. For example, the plateau of NCE scores for the SAT 10 Problem Solving subtest and an insignificant gain in the scale scores for the ARMT Math subtest were seen during the 2005-2006 school year, the year in which fewer mathematics coaches were available to work with the MMI-trained teachers. The SAT 10 Problem Solving subtest and the ARMT Math subtest primarily focus on problem solving skills that are more difficult to teach on the part of the teacher and more difficult to understand on the part of the student. The basis of the MMI is to train teachers to teach for understanding by establishing in teachers a deep understanding of mathematics and by supporting the use of teaching strategies that create a deep understanding of mathematics for students.

The historical overview of mathematics education described in the literature reveals that Chinese teachers have a much deeper understanding of elementary mathematical concepts than their U.S. counterparts do. A goal of the MMI is to ensure that mathematics coaches have this deep level of understanding of mathematics and can transfer that deep understanding to the MMI-trained teachers. Mathematics coaches are able to teach and model higher level teaching strategies through planning meetings with the MMI-trained teachers and through follow-up in the classroom. Mathematics coaches are a tremendous and valuable resource for the MMI-trained teacher, who returns to the classroom with teaching manuals and new manipulatives to incorporate into lesson plans. When the mathematics coaches are removed, the huge support piece sustaining the gains made by the MMI-trained teacher in the knowledge of more effective teaching practices no longer exists. Coaches who have been answering questions, modeling strategies, and offering feedback are no longer available at the necessary level of support; the importance of the newly acquired knowledge and of the experience of teaching mathematics on a deeper level is no longer adequately reinforced.

Implications

The critical-thinking model of focusing on problem solving and on the ability to articulate one's thinking is often denied to high poverty students. The MMI provides this model and constitutes a pedagogy that is atypical for poverty. This model of high expectations could equate with the way gifted children were taught 20 years ago. The MMI is not a pull-out program; it is a method of teaching. In this study of three high poverty schools, the Hispanic and Black students, who comprised 82% of the combined populations of the three schools, as well as the subset of LEP students, made statistically significant gains in achievement scores with the use of this teaching method; in all schools, students of MMI-trained teachers seemed to find success. The implication for mathematics education is that this method of constructivist teaching and learning seems to be effective for all students and would be appropriate for any school working on improving understanding of mathematics; however, the method seems to be effective for high poverty schools.

Schools that choose to utilize effective teaching methods to improve student understanding of mathematics must be prepared to make a substantial investment. Fully implementing the MMI requires a financial commitment for at least 4 years. Clear expectations related to the initiative must be met. There must be district commitment and curriculum alignment. There are expectations for strong principal leadership, low teacher turnover, coaching support, training, materials, follow-up, and leadership training. When discussing the MMI with its director, Susan Pruet, the three most important elements for success that emerge after teacher training are, in order of importance, principal leadership, teacher turnover, and mathematics coach involvement. Pruet reported that, in Mobile, Alabama, the two schools that had those elements in place when the initiative began emerged as the achievement leaders among the MMI schools. As the other MMI-trained schools were able to get all three elements in place, the gain in scores equalized over the MMI schools in approximately 5 years.

In this study, the researcher noted that many factors could account for the differences in the scores of the study schools; however, additional variables would be needed to more fully understand the differences. School A started out with a substantially lower mean NCE score for the SAT 10 Procedures and Problem Solving subtests than the other schools did but was able to more than double the mean score for the SAT 10 Procedures subtest from baseline to the final testing occasion (NCE scores = 26.63 and 53.77, respectively) and to substantially increase the mean score for the SAT 10 Problem Solving subtest (NCE scores = 30.73 and 43.98, respectively). School C made gains in NCE scores for the SAT 10 Procedures subtest over the three testing occasions (NCE scores = 44.46, 44.88, and 46.27) and for the SAT 10 Problem Solving subtest (NCE scores = 41.12, 46.79, and 42.47); however, the net gains were not significant and were not as great as those made in the other schools (School A NCE scores for Procedures = 26.63, 44.85, and 53.77; School B NCE scores for Procedures = 45.11, 49.52, and 60.76; School A NCE scores for Problem Solving = 30.73, 39.29, and 43.89; School B NCE scores for Problem Solving = 41.18, 48.08, and 48.37; no data are reported here for the ARMT Math subtest because there were no baseline data available for comparison across all three testing occasions). After thinking this study and the possible reasons School C did not make significant gains, the researcher concluded that investigating the three elements of principal leadership, teacher turnover, and mathematics coach involvement may help to elucidate the findings that cannot be explained at the present time and may prove to have an equalizing effect in the school system used in the study, as such an investigation did in Mobile, Alabama, in Pruet's (2005) study.

It is noted that, at the time of this writing, Dr. Honi Bamberger is the key trainer in the MMI initiative. Dr. Bamberger is known for her expert teaching and integral involvement in the MMI initiative. She relates to teachers in a respectful, intelligent manner and models her methods of teaching and classroom management. Bamberger conducts the 2-week institute in a simulated classroom, and teachers and principals follow her like children followed the Pied Piper. The benefit of her master teaching abilities to the MMI is her positive influence on those she trains; however, because one individual will not sustain an initiative indefinitely, the issues of leadership and sustainability must be addressed. The implication for the present study is that, in any "train-the-trainer" model, the key trainer must be extremely competent and effective.

Last, the benefit over time (between testing occasions) is that there were statistically significant gains made in the SAT 10 Procedures and Problem Solving subtest scores over the 2 years of intervention. Because gains were seen over time (testing occasions) the researcher surmises and anticipates that those gains will continue over additional testing occasions and that there will be a correlation between the number of consecutive years that a student has been taught by an MMI-trained teacher and the students' understanding of mathematics as evidenced by student achievement scores. The implication is that, if the MMI strategies used in elementary school mathematics instruction continue to be used in middle school mathematics instruction, there will be a continued deepening of mathematics understanding despite the concepts' becoming would become more sophisticated. John Mayer (as cited in Powell & Shuler, 2006), associate chair and professor of mathematics at the University of Alabama at Birmingham, recognizes the need "for more specialized training for people who teach math at the middle-school level" (p. 1). Mayer (as cited in Powell & Shuler) is involved in teacher-training efforts at the middle-school level that are based on the concept of

> helping teachers construct practical exercises that will allow their students to discover basic math concepts instead of just being told about them. The idea is to show students how to use math as a tool for solving real-world problems, which in turn should spark their curiosity and make them want to learn more. . . . The new classes are being designed to give future teachers a deeper understanding of the material they will be teaching in middle school. That way, they will be able to explain their reasoning algebraically and geometrically. (pp. 1-2)

Conclusions

In analyzing the study data for each of the null hypotheses, the researcher noted a statistical significance for the interaction between time (testing occasions) and schools. Additionally, there was a significant interaction between mean NCE scores for the SAT

10 Procedures and Problem Solving subtests and time (testing occasions). Important results from the study are that two findings were consistently identified:

1. Findings for Null Hypotheses 2 and 3 showed that LEP students consistently outperformed non-LEP students on the ARMT Math subtest and the SAT 10 Problem Solving subtest after 2 years of intervention by MMI-trained teachers.
2. Findings for Null Hypotheses 1, 2, and 3 showed that Black and Hispanic students outperformed White students on the ARMT Math subtest and the SAT 10 Procedures and Problem Solving subtests after 2 years of intervention by MMI-trained teachers.

Recommendations for Improving the Research

For the purposes of this study, it was important to have an established baseline for measuring gains caused by the MMI teacher training. This study utilized data from the three high poverty schools that chose to give the SAT 10 in the second grade. These baseline data were provided through this second-grade SAT 10 testing that is not required by the school system or the Alabama Department of Education. Only three MMI schools conducted second-grade SAT 10 testing; therefore, the number of students and schools to study was limited. The sample sizes for the independent variables were too small in many of the cells to detect a statistically significant interaction (if one existed) by using ANOVA with repeated measures; therefore, paired and independent *t* tests were necessary to explore significant changes in the data over testing occasions. The research findings possibly would have been more robust if there had been more students and schools to include in the data analysis so that cell sizes would have been large enough to utilize ANOVA with repeated measures. Perhaps the study design could have included schools from other school districts to increase sample sizes in the independent variable sub-groups.

Another improvement would be to control for other variables that could have affected the results of the study, such as the variables of (a) amount of time the MMI-trained teachers spend with mathematics coaches, (b) level of implementation of the MMI (i.e., teacher implementation, principal commitment), and (c) rate of MMI-trained teacher turnover. First, the mathematics coach is a critical and vital component of the MMI. Although the researcher did not anticipate a change in the ratio of mathematics coaches to MMI-trained teachers during the study, a change took place in this variable that could have significantly affected the results of the study. Second, the level of implementation of the MMI could have been different for the study schools; recommendations to improve the research include defining levels of implementation of the MMI program elements and including these in analyses. Last, including MMI-trained teacher turnover as a factor in future research is important.

Recommendations for Future Research

Because schools that are not considered high poverty also participate in the MMI, future research could focus on those schools, as well. The current study was intentionally designed to study high poverty schools because the MMI originated to serve this population. However, researching the impact of the MMI on all students will potentially broaden the usefulness of the MMI to other schools that may choose to participate and will expand the generalizability to students in other socioeconomic groups.

Another potentially beneficial study would be one that compares the achievement of students whose teacher has been trained in the MMI and the Alabama Math, Science and Technology Initiative (AMSTI) with that of students whose teacher has been trained in only the MMI or the AMSTI. The AMSTI is relatively new to Alabama schools at the time of this writing and follows the same philosophy that the MMI does. In some schools, MMI-trained teachers are also being trained in the AMSTI. The proposed study would help school systems that have participated in the MMI training decide whether training teachers in both initiatives would further improve student achievement.

Future research focusing on the impact of the mathematics coach on student achievement is recommended. As stated previously, there were four mathematics coaches who served 57 MMI-trained teachers in eight MMI schools in 2004-2005. Total time spent by mathematics coaches in each MMI school in the school system averaged 8 hr per week. In 2005-2006, funding issues caused the number of mathematics coaches to be decreased to two coaches serving 109 MMI-trained teachers. Because of the change in the coach-to-teacher ratio, the total time spent by mathematics coaches in each MMI school dropped to 8 hr every other week, with less time spent per teacher. Because support and accountability for the MMI-trained teachers by the mathematics coaches is an integral aspect of the MMI, the shift in the ratio of coaches to teachers and the corresponding reduction in the rate of increase in student achievement caused the researcher to conclude that a subsequent study including this variable is necessary and important. At the time of this writing, the school system has increased the staff of mathematics coaches to four mathematics coaches serving 140 MMI-trained teachers. A carefully designed study that captures the amount of time that mathematics coaches expend in support of each MMI-trained teacher would facilitate research into the effect of this variable and provide more detailed examination of this obviously successful learning initiative that has helped high poverty students learn mathematics in the schools incorporated into the current study.

In conclusion, the study findings lead the researcher to believe that the MMI has a positive impact on student achievement in high poverty schools. Furthermore, the MMI warrants continued investigation to more fully explore its potential to impact student achievement in the 21st century.

LIST OF REFERENCES

*A Nation at Risk. (*1983). Retrieved September 30, 2006, from http://www.ed.gov/pubs/NatAtrisk/risk.htm

Alabama Reading and Mathematics Test Technical Manual. (2005). San Antonio, TX: Harcourt Assessment.

American Institutes for Research. (2006, January). *New study of the literacy of college students finds some are graduating with only basic skills*. Washington, DC: Author.

Association of American Colleges and Universities, Greater Expectations National Panel. (2002). *Greater expectations: A new vision for learning as a nation goes to college.* Washington, DC: Author.

Ball, D. L. (2003, February). *What mathematical knowledge is needed for teaching?* Paper presented at the 2003 University of Michigan Secretary's Summit on Mathematics, Washington, DC. Retrieved June 24, 2005, from http://www.ed.gov/rschstat/research/progs/mathscience/ball.html

Ball, D. L., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433-456). New York: Macmillan.

Bamberger, H. (Ed.). (2005). *Mobile Math Initiative: Teaching mathematics for meaning and mastery summer institute participant's notebook*. Mobile, AL: Mobile Area Education Foundation.

Catholic School Department. (n.d.). *Mathematics standards glossary*. Retrieved September 30, 2006, from http://www.csdsac.org/Curriculum/mathglossary.htm

Coeyman, M. (2003). Twenty years after "A Nation at Risk" [Electronic Version]. *Christian Science Monitor, 95,* 13. Retrieved September 30, 2006, from http.//www.C.S.Monitor.com/2003/04/22/p13so2-lepr.html

Cohen, D. K. (1989). Teaching practice: Plus change. In P.W. Jackson, (Ed.), *Contributing to educational change: Perspectives on research and practice* (pp. 27-84). Berkeley, CA: McCutchan.

Cole, R.W. (Ed.). (1995). *Educating everybody's children: Diverse teaching strategies for diverse learners. What research and practice say about improving student achievement.* Alexandria,VA: Association for Curriculum and Development.

Covey, S. (1989). *The 7 habits of highly effective people*. New York: Simon and Schuster.

Creswell, J. W. (2003). *Research design.* Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Pearson.

Darling-Hammond, L., & Sykes, G. (Eds.). (1999). *Teaching as the learning profession: Handbook of policy and practice*. San Francisco: Jossey-Bass.

Davis, C. S. (2002). *Statistical methods for the analysis of repeated measures*. New York: Springer.

Forgione, P. D., (1998). *Education statistics and education policy: The American experience*. Retrieved June 24, 2007, from the U.S. Department of Education, National Center for Education Statistics Web site: http://nces.ed.gov/Pressrelease/amex.asp

Fosnot, C. (1996) Constructivism: A psychological theory of learning. In C. Fosnot (Ed.), *Constructivism: Theory, perspectives and practice* (pp. 8-33). New York: Teachers College Press.

Fosnot, C., & Dolk, M. (2001). *Young mathematicians at work: Constructing multiplication and division*. Portsmouth, NH: Heinemann.

Fosnot, C., & Dolk, M. (2002). *Young mathematicians at work: Constructing fractions, decimals, and percent*. Portsmouth, NH: Heinemann.

Green, S, B., Salkind, N. J., & Akey, T. M. (1999). *Using SPSS for windows: Analyzing and understanding data* (2nd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Gollnick, D., & Chinn, P. (2002). *Multicultural education in a pluralistic society* (6^{th} ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Hiebert, J. (2003). What research says about the NCTM standards. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 5-23). Reston, VA: The National Council of Teachers of Mathematics.

Kamii, C. (1989). Young children reinvent arithmetic. New York: Teachers College Press.

Kilpatrick, J. (2001). Where's the evidence? *Journal for Research in Mathematics Education, 32*, 421-427.

Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American children.* San Francisco: Jossey-Bass, 1994.

Leinwand, S. (2000). *Sensible mathematics: A guide for school leaders*. Portsmouth NH: Heinemann.

Lemke, M., A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., & Jocelyn, L. (2004). *International outcomes of learning in mathematics literacy and problem solving: PISA 2003 results from the U.S. perspective* (NCES Publication No. 2005-003). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Ma, L. (1999). *Knowing and teaching elementary mathematics: teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Erlbaum.

Mertler, C. A. & Vannatta, R. A. (2001). *Advanced and multivariate statistical methods*. Los Angeles: Pyrczak Publishing.

McGuire, K. (2000). *Ed. Dept. names its pick for math programs*. Retrieved December 28, 2005, from http://www.nctm.org/news/articles/1999-12programs.htm

National Council of Teachers of Mathematics. (2000). *Principals and standards for school mathematics*. Reston, VA: Author.

National Science Foundation. (1996). *Shaping the future*. Retrieved November 1, 2005, from http://www.pkal.org/documents/ShapingTheFuture1996Page11.cmf

National Staff Development Council. (2001). *NSDC standards for staff development*. Oxford, Ohio: Author.

Nussbaum-Beach, S. (2005). *Powerful conversations*. Retrieved January 14, 2007, from http://21stcenturylearning.typepad.com/alabama_best_practices_ce/

No Child Left Behind Act, PL 107-110 (2001).

Organization for Economic Cooperation and Development. (2003). *Program for international student assessment*. Washington, DC: National Center for Education Statistics.

Osborne, J. W., & Overbay, A. (2004). *The power of outliers*. Retrieved June 17, 2007, from http://pareonline.net/getvn.asp?v=9&n=6

Parker, R. (2005). *Number talks*. Retrieved October 10, 2006, from http://web001.greece.k12.ny.us/files/filesystem/Number%20Talks.doc

Payne, R. (2001). *A framework for understanding poverty*. Highlands, TX: Aha! Process.

Peske, H., & Haycock, K. (2006). *Teaching inequality: How poor and minority students are shortchanged on teacher quality*. Washington, DC: The Education Trust.

Phipps, R. (1998). *College remediation, what it is, what it costs, what's at stake.* Washington, DC: Institute for Higher Education Policy.

Powell, P., & Shuler, R. (Eds.). (2006). Creating a new kind of math teacher. *UAB Education, 8*(2), 1-3.

Pruet, S. (2005). The Maysville to Mobile mathematics initiative from the director's perspective. In H. Bamberger (Ed.), *Mobile Math Initiative: Teaching mathematics for meaning and mastery summer institute participant's notebook* (p. 1). Mobile, AL: Mobile Area Education Foundation.

Schoenfeld, A. (2002). Making mathematics work for all children: issues of standards, testing, and equity. *Educational Researcher, 31*(1), 13-25.

Silver, E. (1998). *Improving mathematics in middle schools: Lessons from TIMSS and related research*. Washington, DC: U.S. Department of Education.

Smith, M. S. (2001). *Practiced-based professional development for teachers*. Reston,VA: National Council of Teachers of Mathematics.

Stanford Spring Technical Data Report. (2003). San Antonio, TX: Harcourt Educational Measurements.

Steen, L. (1999). *Twenty questions about mathematical reasoning*. Retrieved December 16, 2005, from the St. Olaf College Web site at http://www.stolaf.edu/people/steen/Papers/reason.html

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematical instruction: An overview of the TIMSS video study. *Phi Delta Kappan, 79*(1), 14-21.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap.* New York: Free Press.

Stigler, J. W., & Hiebert, J. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS video study. *The Elementary School Journal, 101*(1), 7.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*, 12-17.

The Diversity Connection (n.d.). *Glossary.* Retrieved May 21, 2007, from http://ftad.osu.edu/CSP/glossary.html

U.S. Department of Education. (2002). *The condition of education 2002*. Retrieved August 11, 2003, from U.S. Department of Education, Institute of Educational Sciences, National Center for Educational Statistics Web site: http://nces.ed.gov/pubs2002/2002025.pdf

U.S. Department of Education. (2003). *National assessment of educational progress: The nation's report card, 2003.* Retrieved September 30, 2006, from http://nces.ed.gov/nationsreport card

U.S. Department of Education (n.d.). *Building the legacy: IDEA 2004.* Retrieved June 2, 2007, from http://idea.ed.gov/explore/view/p/%2Croot%2Cstatute%2CI%2CA%2C602%2C29%2C

University of New York. (1999). *Remedial education: New York and the nation.* Retrieved September 30, 2006, from University of New York, State Education Department, Office of Higher Education Web site: http://www.highered.nysed.gov/oris/remedrpt.htm

Van Hanegan, J., Pruet S., & Bamberger, H. (2004). Mathematics reform in a minority community: student outcomes. *Journal of Education for Students Placed at Risk, 9*, 189-221.

White, S., & McCloskey, M. (2005). *Framework for the 2003 assessment of adult literacy.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

APPENDIX A

THE MOBILE MATH INITIATIVE

Teaching the Mobile Math Initiative Way is often described as "teaching all children mathematics for meaning and mastery so they can make sense of the mathematics they do, and thus 'own their mathematics'" (Pruet, 2005, p. 1). According to the director of the Mobile Math Initiative, Dr. Susan Pruet, the initiative

> is an intensive, comprehensive and year round professional development program for teachers and administrators that transforms the quality of mathematics instruction, the school climate, and culture in order to raise student achievement. . . . The focus of Mobile Math Initiative is student learning; therefore, the teachers' professional development is considered to be extremely important because there is a direct correlation between high levels of teacher quality and increased student achievement. (p. 1)

The Mobile Area Education Foundation and the Mobile County public school system formed a partnership with the goal of improving students' mathematics achievement, which in turn would help to raise the graduation rate. The development of the Mobile Math Initiative began in 1998; after several years of planning, seven key components evolved: (a) school-wide commitment, (b) professional development and support, (c) curriculum and assessment, (d) family involvement, (e) leadership, (f) school climate, and (g) partnerships. The initiative was started Maysville, a lower socioeconomic area of Mobile with four elementary schools, two middle schools, and one high school. The initiative is being replicated in other schools in Mobile, as well as in the Hoover school system and the targeted school district.

When planning the initiative, Dr. Pruet saw a need to hire a professional developer for the initiative. Dr. Honi Bamberger was hired to fill this position. She is a former elementary mathematics resource teacher and college professor. Dr. Bamberger served as the co-principal investigator of Project IMPACT, a successful mathematics initiative in the 1990s with goals that served as a foundation for the MMI. Project IMPACT focused on high poverty schools. Dr. Bamberger also coordinates teams of mathematics professionals who assist school systems in implementing mathematics reform.

APPENDIX B

INSTITUTIONAL REVIEW BOARD FOR HUMAN USE (IRB) APPROVAL

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Institutional Review Board for Human Use

Form 4: IRB Approval Form
Identification and Certification of Research
Projects Involving Human Subjects

UAB's Institutional Review Boards for Human Use (IRBs) have an approved Federalwide Assurance with the Office for Human Research Protections (OHRP). The UAB IRBs are also in compliance with 21 CFR Parts 50 and 56 and ICH GCP Guidelines. The Assurance became effective on November 24, 2003 and expires on February 14, 2009. The Assurance number is FWA00005960.

Principal Investigator: PAYNE, JEANNE

Co-Investigator(s):

Protocol Number: **X070110002**

Protocol Title: *An Investigation of the Impact on Student Achievement of the Mobile Math Initiative in Three High Poverty Schools in a Single School District*

The IRB reviewed and approved the above named project on 02/06/07. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services. This Project will be subject to Annual continuing review as provided in that Assurance.

This project received EXPEDITED review.

IRB Approval Date: 2-6-07

Date IRB Approval Issued: 02/06/07

Marilyn Doss

Marilyn Doss, M.A.
Vice Chair of the Institutional Review Board for Human Use (IRB)

Investigators please note:

The IRB approved consent form used in the study must contain the IRB approval date and expiration date.

IRB approval is given for one year unless otherwise noted. For projects subject to annual review research activities may not continue past the one year anniversary of the IRB approval date.

Any modifications in the study methodology, protocol and/or consent form must be submitted for review and approval to the IRB prior to implementation.

Adverse Events and/or unanticipated risks to subjects or others at UAB or other participating institutions must be reported promptly to the IRB.

470 Administration Building
701 20th Street South
205.934.3789
Fax 205.934.1301
irb@uab.edu

The University of
Alabama at Birmingham
Mailing Address:
AB 470
1530 3RD AVE S
BIRMINGHAM AL 35294-0104