EXAMINING THE RELATIONSHIP BETWEEN PARTICIPATION IN A MATH SCIENCE PARTNERSHIP AND CHANGES IN STUDENT OUTCOMES IN HIGH SCHOOL MATHEMATICS USING ACTIVITY THEORY AS A LENS

A Dissertation
Submitted to
The Temple University Graduate Board

in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

by
Kathleen Krier
May 2011

Dissertation Examining Committee:
Jacqueline Leonard, Committee Chair, CITE/Math & Science
Catherine Schifter, CITE/Educational Administration
James Earl Davis, Interim Dean, College of Education
Joseph P. DuCette, Psychological Studies in Education, Chairperson
Will J. Jordan, Educational Leadership and Policy Studies

ABSTRACT

EXAMINING THE RELATIONSHIP BETWEEN PARTICIPATION IN A MATH SCIENCE PARTNERSHIP AND STUDENT OUTCOMES IN HIGH SCHOOL MATHEMATICS USING ACTIVITY THEORY AS A LENS

By Kathleen Krier

Doctor of Education

Temple University, May, 2011

Major Advisor: Dr. Jacqueline Leonard

Using Activity Theory as both the theoretical and analytical framework, this study investigated whether participation in the Math Science Partnership of Greater Philadelphia (MSPGP) had any relationship with changes in PSSA scores and math course taking patterns for students attending 23 partner district high schools. Participation included time spent attending professional development provided by the MSPGP and the level of engagement that a district had with the partnership. Results showed significant gains in PSSA achievement overall ($t = 4.03$, $df = 22$, $p = .001$) and for African American students in particular ($t = 2.53$, $df = 13$, $p = .025$). These results are similar to statewide results in Pennsylvania. The biserial correlation showed a relationship between the overall improvement in PSSA scores and the level of professional development in which schools participated ($r_p = .62$, $p < .05$). Results also showed a significant gain in higher level math course completions for Latino/a students ($t = 3.08$, $df = 19$, $p = .006$). Although the gain in PSSA scores was not significant for Latino/a students, it was strongly correlated with the level of engagement schools had with the MSPGP ($r_p = .90$, $p < .05$). Course completions, both overall and higher level, increased for the majority of schools. Achievement gaps persist for many of the schools even though the majority of African American and Latino/a students had increases in PSSA achievement as well as increases in course completions. The study also investigated whether other independent school district variables had any relationship with changes in outcomes. Results of the Pearson Correlation showed a significant relationship between the use of reformed curriculum and changes in PSSA achievement (r = .43, p = .04).

This study was completed utilizing data collected by the Math Science Partnership of Greater Philadelphia. The author of this study was an employee of the MSP; however, no funding from the National Science Foundation was provided for the completion of this study. The findings and conclusions are solely those of the author and do not necessarily reflect the views of the Math Science Partnership of Greater Philadelphia or the National Science Foundation.

ACKNOWLEDGEMENTS

Completing this study as well as the doctoral program at Temple University required years of commitment on my part. However, I would not have been able to make such a commitment without the help, support, and love of many individuals in my life. First, I would like to thank Dr. Jacqueline Leonard. Words cannot adequately express my gratitude. Dr. Leonard always provided feedback and advice that helped me immensely during my time at Temple. She is an invaluable mentor who taught me much about urban education in the city I love. She provided a guiding hand as I learned how to navigate through the research process as well as a sympathetic ear as I dealt with the many obstacles that I faced while on this journey. Her encouragement always lifted me up when I felt defeated. I will be forever grateful that I had the opportunity to work with and learn from Dr. Leonard.

Many thanks to Dr. Schifter, whose input was critical in the beginning stages of writing my proposal and in revising chapter three so that my methodology was clearly and thoughtfully stated. Thanks to Dr. DuCette, for helping me with the statistics and to Dr. Davis and Dr. Jordan, who graciously accepted my invitation to participate on my committee.

My final thanks to the Temple community goes to Dr. Steve Ryan, who had the vision to help me organize a large amount of data into a coherent study. He was able to see a story in the data; something that I had great difficulty with because I was so entrenched in it. His patience as I ruminated about a direction for my research allowed me to grow professionally as a researcher.

I owe many thanks to: Deanna, Barbara, Nancy, and Deborah, members of the MSPGP family, whose institutional memories were essential in my ability to organize and interpret the data.

Reaching this milestone would not be possible without the support of my family. I would like to thank my daughters, Marissa, Sara, and Beth, who unselfishly accepted the time commitment I needed to make and understood that time devoted to research and writing was time away from them. I look forward to spending time with them again and with my granddaughter, Lila.

Finally, I would like to thank my husband, Steve. Without his love, understanding, and enduring support over the last several years, I would not have accomplished all that I have. He lovingly sacrificed time with me and took on many responsibilities to allow me to reach my goal.

TABLE OF CONTENTS

PAGE

PAGE

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1
INTRODUCTION

Purpose

It has been over 56 years since the landmark court decision of Brown versus the Board of Education of Topeka. Since then, countless reform efforts have been attempted in order to provide equal opportunities for all students. During the 1980s and 1990s the National Council of Teachers of Mathematics (NCTM) developed rigorous standards for the teaching and learning of mathematics. NCTM identified equity as its first principle and “demands that high expectations for mathematics learning be communicated in words and deeds to all students” (NCTM, 2000, p. 11). Despite this recommendation and demand, inequity in education, particularly in mathematics and the sciences persists.

Although there was some improvement in mathematics achievement during the 1970s and 1980s, when the results of the National Assessment of Educational Progress (NAEP) showed increases in achievement for minority students and the narrowing of achievement gaps, the trend has reversed, with achievement gaps creeping up (Lee, 2002). The gaps that exist today are similar to what they were at the beginning of the 1980s (Lee, 2004). Results for the 2009 NAEP showed that 13% of African American and 19% of Latino/a 8th-graders were proficient, compared to 55% of their white peers (NCES, 2010). This is higher than the results for 2007, when 9% of African American and 13% of Latino/a students were proficient (Flores, 2004).

In Pennsylvania, during 2009, 28.3% of 11th grade African American students and 30.7% of 11th grade Latino/a students scored proficient or better on the Pennsylvania Statewide System of Assessment (PSSA) compared to 61.6% of white students (PA Department of Education, 2009). National SAT results show disparities as well. Although the largest number of minorities in the history of the SAT took the test during 2009 (40% of all test takers), African American and Latino/a students' mean scores were well below that of Whites, by 85 and 110 points, respectively (College Entrance Examination Board, 2009).

Moreover, achievement in mathematics in the United States, in general, pales compared to that of many other nations. During 1987, the results of the Second International Mathematics Study (SIMS) were made public. Findings for U.S. 8th-grade students revealed that students' scores were similar to the international average for procedural math but far below it in challenging, non-routine mathematics (McKnight et al., 1987). On the high school level, the United States scored in the bottom one-fourth in: algebra, geometry, and elementary functions. Two decades later, the results of the fourth administration of the Trends in International Math and Science Study (TIMSS) show improvement for both fourth and eighth grades, with fourth grade ranking eleventh out of 36, and eighth grade ranking ninth out of 48 (Gonzales, Williams, Roey, Kastberg, & Brenwald, 2008). While these results are promising, high school students were not evaluated in this study; therefore, one cannot say whether or not improvements at the secondary level occurred.

In an effort to remedy the chronic problem of low academic achievement in public education, the federal government enacted Public Law PL 107-110, the No Child Left Behind Act (NCLB) during 2001, which reauthorized the Elementary and Secondary Education Act of 1965. The intent was to boost achievement overall and specifically target underrepresented groups to eliminate achievement gaps. Shortly after the passage of NCLB, Representative Sherwood Boehlert of New York and House Science

Committee Chair, introduced The Math and Science Education Partnership Act, H.R. 1858 (Mervis, 2001). Math and Science Partnership "projects are expected to raise the achievement levels of all students and significantly reduce achievement gaps in the mathematics and science performance of diverse student populations" (NSF, 2002, n. p.).

The purpose of the current study is to ascertain whether a high school's participation in the Math Science Partnership of Greater Philadelphia (MSPGP) had any effect on student outcomes in mathematics, specifically at the secondary level. The intent of the study is to contribute to the growing body of literature examining student achievement and inequity in K-12 mathematics education. Expanding this body of knowledge is important because increasing mathematical literacy and closing achievement gaps helps to build human capital for the nation (Checchi, 2006; Nettles & Millett, 1999). Mathematical literacy is positively correlated with the completion of post-secondary education (Adelman, 2006) and earnings (Rose & Betts, 2004). Moreover, levels of educational attainment are directly correlated with economic earning power and upward mobility, with those holding a Bachelor's degree earning 60.4% more than those with only a high school diploma (Liming & Wolf, 2008). According to Fullan (2007):

> The global society is increasingly complex, requiring educated citizens who can learn continuously, and who can work with diversity locally and internationally. Although the source of blame varies, it is now an undeniable conclusion that the education system and its partners have failed to produce citizens who can contribute to and benefit from a world that offers enormous opportunity and equally complex difficulty finding one's way in it. (p. 7)

Therefore, continuing the line of inquiry investigating student achievement and equity is warranted.

For the purpose of this study, student outcomes in high school mathematics are defined as: achievement levels on the mathematics portion of the 11th-grade PSSA and the percentage of students completing high school level college preparatory mathematics courses. The study examines both overall course completion and higher level course completion. It also investigated whether there were any changes in achievement gaps between white and minority students. The unit of analysis for this study is the high school, given that MSPGP project data were not gathered at the individual student level. However, since the data are disaggregated by race and grade level, comparisons between subpopulations within each high school will be made in order to determine whether improvements, if they occurred, were equitable.

Background

In order to understand the underpinnings of the development of the Math Science Partnership program, a brief historical context is presented here. The National Science Foundation (NSF) was established on May 10, 1950 through the passage of Public Law 81-5007. Its purpose was to advance scientific research, development, and innovation as a matter of national security during the years following World War II. Thus, NSF played a vital role in science education reform from its inception (Mazuzan, 1994).

The launch of Sputnik during 1957 spawned a national urgency to focus on mathematics and science education. It was vital to grow enough new scientists to meet the challenges of: the space race; global competiveness; and, maintaining national security (Waterman, 1960). Meanwhile, beginning in 1953, NSF was immersed in a concerted effort to better prepare science teachers through its summer institutes program for college science teachers. The program expanded to include high school teachers during 1954 (Wolfle, 1957) and elementary teachers during 1957 (Wailes, 1968). Year-long institutes were added during 1957 for high school teachers and during 1959 for elementary teachers (Waterman, 1960). They focused on science and mathematics content, pedagogy, and laboratory equipment use. The institute program for high school teachers ended abruptly during 1982 when the Reagan administration cut the budget for the NSF eliminating all K-12 education programs (Brazeel, 1981).

Improving curricula was also a priority for NSF. The Course Content Improvement Program was established during 1957. Its purpose was to bring math and science course content material up to date (Lomask, 1976). The program primarily developed challenging materials that targeted college-bound students. By 1970, NSF realized that advancing the scientific frontier required more experts than were currently enrolled in post-secondary education and the focus shifted to developing curricula that would promote science and mathematics literacy for all students (Lollor, 2000). This mission continues today through the MSP program.

During 1983, the sobering report, *A Nation at Risk* was released. It had been over 25 years since the launch of Sputnik, yet the National Commission on Excellence in Education had determined that among other deficits in the American system of education the "average achievement of high school students on most standardized tests is now lower than 26 years ago when Sputnik was launched" (National Commission on Excellence in Education, 1983). Despite NSF's efforts to improve achievement, students seemed to be no more scientifically or mathematically literate than they were two decades prior.

By 1991 the NSF was poised to expand its mission even further through the creation of the Statewide Systemic Initiatives program (SSI). The SSI program was a leap of faith for NSF; it had little experience in supporting large-scale restructuring efforts (Mervis, 1998). “The SSIs were intended to create substantial and enduring reform in education policy, administration and management, support services, and the teaching and learning of mathematics and science” (Heck & Weiss, 2005, p. 3). The overarching goal of the SSI program was grounded in equity. All SSI activities would be in accordance with the NSF philosophy that all children can learn high quality mathematics and science (Zucker, Shields, Adelman, Corcoran, & Goertz, 1998). This strong belief figured prominently in the development of the MSP program. Eventually, the program expanded to include Rural Systemic Initiatives (RSI), Urban Systemic Initiatives (USI), and Local Systemic Change through Teacher Enhancement (LSC) programs during 1993 through 1995 as NSF began to support smaller systemic changes, such as those initiated by cities and school districts (McKeown, Klein, & Patterson, 2001; Mervis, 1998).

The NSF Statewide Systemic Initiative Program ended in 1998. The concept of an across-the-board large scale systemic initiative never completely came to fruition. There were too many factors associated with reform that were outside the purview of the NSF, such as: teacher certification requirements, the length of the school day, pedagogical practices, union agreements, and the like. Nevertheless, some good did come out of the lessons learned from the program that continues today, most notably, the continued effort for the alignment of teaching, learning, and assessment with a set of rigorous standards. Furthermore, the program brought to light an understanding that building capacity is a complex task that takes longer than five years. Though the vision of the SSI program was never full realized, the lessons learned through its lack of success may have influenced the developers of the MSP program to focus on capacity building as a necessary precursor to improving achievement.

During September 2002, the National Science Foundation (NSF) and the Department of Education (DOE) announced 24 awards under a new Math and Science Partnership program. Together they provided $240 million over five years to educational affiliations across 11 states (NSF, 2004). In the first round of solicitations the NSF articulated the following goals:

> The MSP program seeks to improve K-12 student achievement through a sharp focus on three inter-related issues:
>
> - Ensuring that all students have access to, are prepared for and are encouraged to participate and succeed in challenging and advanced mathematics and science courses;
> - Enhancing the quality, quantity and diversity of the K-12 mathematics and science teacher workforce; and
> - Developing evidence-based outcomes that contribute to our understanding of how students effectively learn mathematics and science. (NSF, 2003, n. p.)

Given the goals of the program, the MSP provides the ideal setting to investigate whether achievement and equity in science and mathematics education have improved for the K-12 schools that were partners in the MSP program.

Statement of the Problem

It has been nine years since the passage of NCLB. The Math Science Partnership program was launched a year later as a policy effort to combat the chronic low performance of students in mathematics and the sciences. Research and evaluation on the MSP program is beginning to emerge in the literature, but understandably, remains sparse. A keyword search of the EBSCOhost databases found 49 scholarly articles containing the words “math and science partnership” A preliminary review of the literature revealed two chief levels of discussion: establishing a set of evaluative models for investigations and research protocols on MSPs (Scherer, 2008; Wong & Socha, 2008; Yin, Hackett, & Chubin, 2008) and, teachers and teaching (see for example, Firestone, Mangin & Martinez, 2007; Firestone, Mangin, Martinez, & Polovsky, 2005; Fisler & Firestone, 2006; Horn, 2008; Moyer-Packenham, Bolyard, & Kitsantas, 2008; Tyler, & Vitanova, 2008). Since the MSP program is less than a decade old, and given NSF’s declaration that MSPs are “research and development efforts” (NSF, 2003, n. p.), it is critical to establish rigorous standards for research and evaluation and to understand the impact the program may have on teachers’ instructional practices and students’ achievement levels. As a reform effort for mathematics and science education, MSPs provide vast opportunities to investigate research questions focused on student achievement. In light of No Child Left Behind (2001), this is especially important in the area of student outcomes in mathematics.

Need for the Study

The MSP portfolio is one that provides several avenues for project development (Scherer, 2008). Research and evaluation efforts will depend on the focus and mission of each project. The focus of one MSP may be on one grade level in a small geographic area, while another, such as the Math Science Partnership of Greater Philadelphia (MSPGP), may have a broad reach across two states, many school districts, and several grade levels. Some MSPs are involved in science education only, while others in mathematics education only. The Math Science Partnership of Greater Philadelphia was involved in both. This study is needed to contribute to the growing body of knowledge on NSF funded MSPs, examining whether and to what extent they impact student achievement.

Theoretical Perspective

The theoretical perspective that serves as a lens for this study is Activity Theory (see Figure 1.1). Activity Theory grew out of the scholarship of Vygotsky (1896-1934), and was further developed by Luria (1976) and Leontj'ev (1978) after Vygotsky's death. This new theoretical perspective went beyond the dominant schools of psychological thought during the 20th-century: psychoanalysis and behaviorism. Its "orientation was a model of artifact-mediated and object-oriented action" (Vygotsky, 1978, p. 40). "The basic activity system comprises a human individual (subject) who interacts (activity) with his/her environment to achieve a goal (object)" (Sanders, 2005, p. 192).

Figure 1.1. Activity Theory: Basic Model

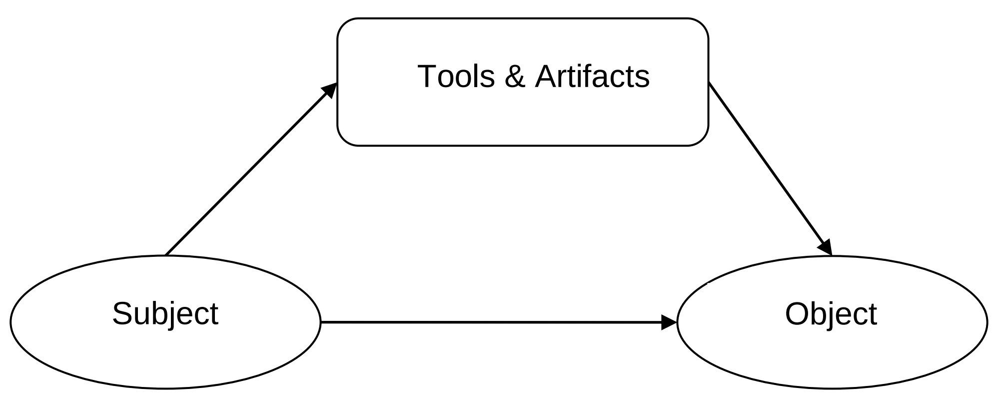


Figure 1.1. Activity Theory: Basic Model Sanders, J. (2005)

This basic model is considered the first generation of Activity Theory. Since its early development, Activity Theory has evolved into second and third generations. Second generation Activity Theory was articulated by Engestrom (1987). While an activity consists of the individual interacting with his/her environment, it is hindered or facilitated by mediating factors. One such factor is the community in which the individual is embedded. Another is the peripheral community with which the individual interacts. A community is shaped by the roles and rules which it defines and refines over time. Roles include the division of labor, both horizontal and vertical, as well as the tasks that are pursued and accomplished. Rules may be implicit or explicit (see Figure 1.2).

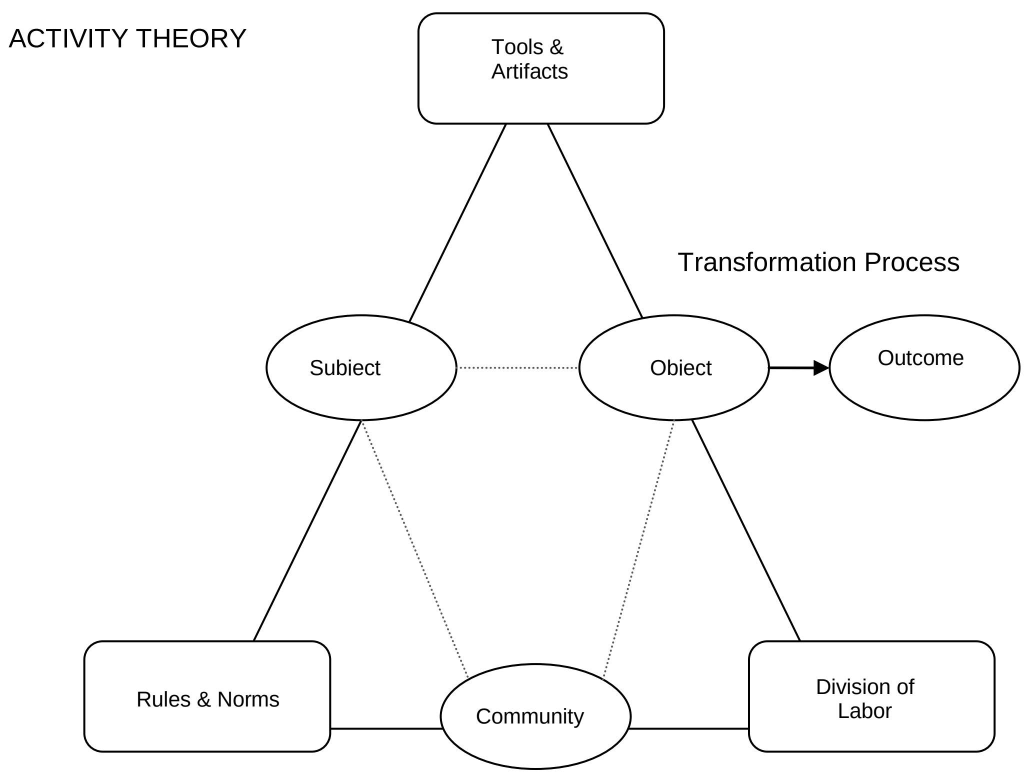


Figure 1.2. Activity Theory: Expanded Model Kuutti, K. (1995)

Third generation Activity Theory examines the interface of two or more dialectical activity systems. In this model boundaries emerge, generating contradictions and tensions as well as the potential to share objects (see Figure 1-3).

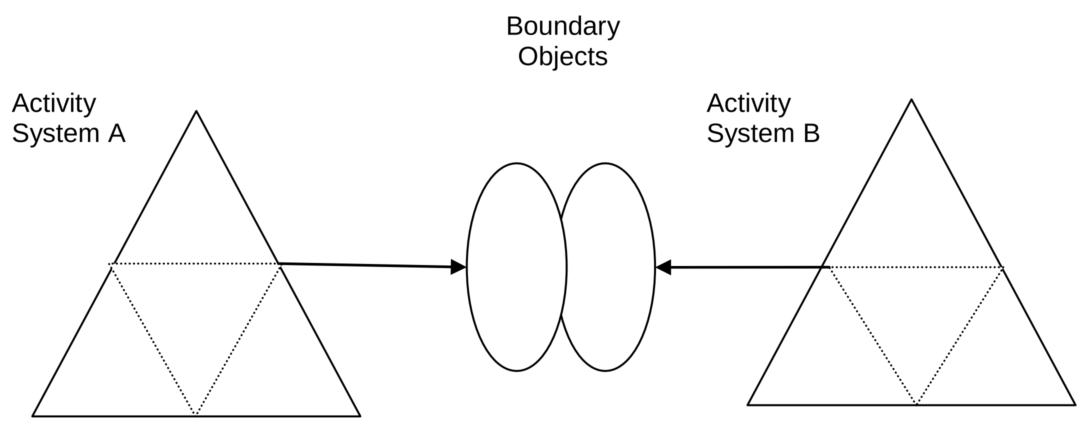


Figure 1.3. Activity Theory: Third Generation Model Engestrom (1987)

The expanded model, or the second generation of Activity Theory, proved useful in guiding the current study because it captures the big picture but also allows one to view the partnership's complexity from several vantage points.

Research Questions

The first goal of the Math Science Partnership program is: "Ensuring that all students have access to, are prepared for, and are encouraged to participate and succeed in challenging and advanced mathematics and science courses" (NSF, 2003, n. p.). Therefore, this study will address the following questions:

1. What relationship, if any, exists between changes in student outcomes, on both the aggregate and disaggregate levels, in secondary mathematics and the level of participation by partner high schools in the Math Science Partnership of Greater Philadelphia?
2. Are there school district factors outside of participation in the Math Science Partnership that may have a relationship to changes in mathematics outcomes for students attending partner high schools?

Summary

In conclusion, addressing these questions is important because the MSP program may prove to be a remedy for chronic underachievement and inequity in science and mathematics education. The next chapter will consult the literature to understand more fully the evolution of the educational partnership and to examine the activities and outcomes that transpire when colleges and universities partner with K-12 educators.

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

The purpose of the current study is to discern whether participation in the Math Science Partnership of Greater Philadelphia (MSPGP) had any relationship with changes in student outcomes in high school mathematics. Math Science Partnerships were mandated by the federal government through the passage of Public Law PL 107-110, the No Child Left Behind Act of 2001 (Title 2, Part B). In order to answer the research questions, the literature was examined on several fronts. Math and Science Partnerships are a relatively new type of educational partnership and the literature is limited. Nevertheless, it will be discussed insofar as it informs the current study. The overarching goal of the MSP program is to “raise the achievement levels of all students and significantly reduce achievement gaps in the mathematics and science performance of diverse student populations” (NSF, 2002, n. p.). The MSPGP offered professional development to teachers and administrators in pursuit of this goal. Thus, the literature on professional development and its relationship to student outcomes will provide one context for the current study. Equally important is the research on student achievement and opportunity to learn; this will provide another context for the study. Furthermore, the notion of partnerships, specifically educational partnerships, will be consulted to bring to bear some challenges faced, lessons learned, and unique features of successful educational partnerships.

Educational Partnerships: The Anecdote for Piecemeal Approaches

For centuries, partnerships have been a way of life for mankind to grow and prosper. Max Weber, 1864-1920, a German sociologist and economist, contends that the emergence of commercial partnerships occurred during the 14th century in Italy in response to the expansion of mercantile trade (Weber, 1905/2003). Weber identified three types of partnerships based on the level of risk assumed and the amount of the capital invested: unilateral, bilateral, and joint household. Over time, the concept of partnership evolved and expanded beyond the commercial partnership. Partnerships are often entered into by individuals or groups who share a common mission or goal, but cannot achieve it alone (Hamos, 2006; Osguthorpe & Patterson, 1998). These types of partnerships often form without monetary benefit and are based on mutuality and trust (Hamos, 2006; Scherer, 2008). K-16 partnerships are often entered into with the mutual goal of improving teaching and learning (Goodlad, 1991). During the last two decades, three types of educational partnerships have emerged:

> primary partner/limited partnerships, which have a managing partner with other organizations providing services to it or to clients; coalition partnerships, [which] boast a division of labor among organizations [and where]…partners are equal but bring different interests and skills to the arrangement; and collaborative [which involve the] division of labor among equal partners; however, decision making is a continuous process shared among partners. (Tushnet, 1993, p. 7)

The concept of educational partnerships emerged during the summer of 1983 in response to *A Nation at Risk* (1983). The presidents of Harvard University, Stanford University, Columbia University, and the Universities of Michigan, Wisconsin, and Chicago, convened the Harvard and Stanford Conference on Public Education, calling for stronger relationships between institutes of higher education (IHEs) and K-12 education (McCormack, 1983). Six ambitious goals emerged from the symposium:

> Improve education through continuing collaboration of university and public school faculty members; Improve teacher education programs; Develop special university programs for enriching the education of precollege students and teachers; Launch research programs aimed at improvement of teaching and learning; Develop university-public school collaborative curriculum development projects; Emphasize teaching's importance by recognizing outstanding teachers. (McCormack, 1983, p. 409)

During 1986, The Task Force on Teaching as a Profession, an arm of the Carnegie Forum on Education and the Economy, issued a report *A Nation Prepared: Teachers for the 21st Century.* It called for bold reforms in teacher education. Its recommendations included: establishing a national board for developing professional teaching standards; obtaining a Bachelor's degree in a discipline as a prerequisite to teacher certification; shifting education courses to the graduate level; recruiting underrepresented groups into teaching careers; and, developing a cadre of teacher leaders.

Simultaneously the Holmes Group, a consortium of deans of colleges of education, began a necessary examination of teacher education, also in response to *A Nation at Risk* (1983). Their work would span more than a decade and three documents would emerge with recommendations for improving: teachers, schools, and schools of education. Their first set of recommendations, *Tomorrow's Teachers* would be released during 1986. Five goals were articulated:

> To make the education of teachers intellectually more solid….To recognize differences in teachers' knowledge, skill, and commitment, in their education, certification, and work….To create standards of entry to the profession—examinations and educational requirements—that are professionally relevant and intellectually defensible….To connect our own

> institutions to schools….[and] To make schools better places for teachers to work, and to learn. (Holmes Partnership, 2006, p. 10)

This would be followed by *Tomorrow's Schools* (1990). One hundred twenty university and college faculty as well as teachers and educators came together to develop the concept of the professional development school. The professional development school is a partnership between colleges/universities and K-12 public schools in an effort to reinvent education. The Holmes Group identified six principles to accomplish this:

> Principle One – Teaching and learning for understanding….Principle Two – Creating a learning community….Principle Three – Teaching and learning for understanding for everybody's children….Principle Four – Continuing learning by teachers, teacher educators, and administrators…. Principle Five – Thoughtful long-term inquiry into teaching and learning…Principle Six – Inventing a new institution. (Holmes Partnership, 2006, p. 98)

The final document in the trilogy, *Tomorrow's Schools of Education* (1995), makes recommendations for a systemic overhaul of colleges of education in order to be successful in developing tomorrow's teachers and tomorrow's schools.

The recommendations that the Holmes group made in their first two documents were echoed by Goodlad (1991). He along with his colleagues formed the Center for Educational Renewal, housed at the University of Washington. The Center conducted a three-pronged investigation to evaluate teacher preparation throughout the nation. First, they conducted "1,800 hours of interviewing college and university presidents, provosts, deans, faculty members, and students" (p. 5) in order to identify and understand how teachers are prepared. The center also investigated the lessons learned in non-educational professions that might lend themselves to improvements in teacher education. Finally, it began the work of growing school-university partnerships for reinventing schools and teacher training. The findings “included the following: (1) a debilitating lack of prestige in the teacher education enterprise, (2) lack of program coherence, (3) separation of theory and practice, and (4) a stifling regulated conformity” (p. 5).

Several years later, the National Commission on Teaching and America’s Future (1996) lamented the same criticisms in their compelling report, *What Matters Most: Teaching for America’s Future*. Upon the completion of a two year study, the Commission found that the following conditions persisted even though the call for reform was well over a decade old:

> Low expectations for student performance; Unenforced standards for teachers; Major flaws in teacher preparation; Painfully slipshod teacher recruitment; Inadequate induction for beginning teachers; Lack of professional development and rewards for knowledge and skill; Schools that are structured for failure rather than success. (pp. 10 & 11)

Among other remedies suggested by the commission was the further development of and commitment to professional development schools, creating partnerships between public schools and the colleges and universities that prepare their teachers.

The Harvard and Stanford Conference on Public Education, the Carnegie Foundation, and the Holmes group brought the notion of school-university partnerships to the forefront of restructuring K-16 education. Since 1983, countless collaboratives between IHEs and public schools have emerged. Two decades ago the U.S. Department of Education estimated that well over 100,000 educational partnerships existed throughout the nation (Rigden, 1991). Twenty years later research around educational partnerships abounds.

Goodlad (1993) likens school-university partnerships to symbiosis. “Symbiosis, in the non-parasitic interpretation of the word means the intimate living together of two dissimilar organisms in a mutually beneficial relationship (p. 29). Although the partners benefit, there are also tensions because of seemingly different cultures (Goodlad, 1993; Kirschner, Dickinson, & Blosser. 1996; Valdez & Snyder, 2006). "Any time unlike partners come together, tensions arise. Because each partner sees things differently, each imagines different solutions to a common problem. This is at once the opportunity and the risk associated with collaborative renewal" (Osguthorpe & Patterson, 1998, p. xix). One theme that resonates throughout the literature examining partnerships is that they thrive and grow as points of contention are dealt with and resolved (Darling-Hammond, 1994; Kirschner, Dickinson & Blosser, 1996; Noffke, Clark, Palmeri-Santiago, Sadler & Shujaa, 1996; Rasch, 1999).

Clifford and Millar (2008) conducted a review of K-20 partnership literature with the following guiding questions: "1. What inquiry methods have been used to study K-20 partnerships and what is the rigor of the methods? 2. How is partnership defined? 3. What do we know about the formation (inputs), process (throughputs), and results (outputs and outcomes) variables associated with a partnership?" (p. 4). Citation tracking software was used to filter 186 studies that investigated educational partnerships; the authors found 36 empirical studies that met their standards for rigor. The criteria by which rigor was determined was: "(a) researcher position (external researchers are considered more rigorous), (b) sampling method, (c) size of samples, and (d) degree to which researchers reported following prescribed and established methods" (p. 6). Twenty-five were single case studies, three were multi-case studies, and eight were survey research. The authors found that most of the studies included convenience samples to gauge successes, and although partnership success is important, challenges and failures, also important, were rarely discussed. The most salient finding of this review was the absence of a definition of partnership. According to the authors, without a precise definition, evaluation of outcomes is problematic. Only 19.4% of the studies reviewed in the analysis met the standards of rigor. This is indicative of the inherent problems in research on educational partnerships.

Professional Development Schools: The Evolution of Educational Partnerships

Professional development schools (PDS) have gained momentum over the last two decades with the belief that this brand of educational partnership may be the solution for comprehensive school reform. “The PDS reform model is compelling because PDS goals assert that educational change should take place simultaneously at K-12 schools and at the university in order for true educational reform to occur” (Campoy, 2002, p. 6). This is analogous to the philosophy espoused by the MSP program in which mathematics and science disciplinary faculty work closely with K-12 educators. Historically, PDSs have provided opportunities for schools and departments of education to provide practical experience for future teachers and for school districts to benefit by having access to a new pool of teachers to draw from, who have knowledge of the school and district culture (NCATE, 2001).

Currently, much of the literature on PDSs examines changes in perceptions, attitudes, and practices of preservice and in-service teachers and sometimes the university faculty who collaborate with them (see for example, Book, 1996; Castle, Fox, & Souder, 2006; Levin & Rock, 2003; Ridley, Hurwitz, Hackett, & Miller, 2005; Swars, Meyers, Mays, & Lack, 2009). Although many PDSs develop with the intent of improving student achievement, few studies have focused on student outcomes (Abdal-Haqq, 1998).

Knight, Wiseman, and Cooner (2000) conducted a study to measure the impact of PDS activities on mathematics and writing. They also examined changes in teacher perceptions. The participants included 284 students in grades two, three, and four. One-third of the students were minorities and 40% were economically disadvantaged. A collaborative teacher research team explored the following question in mathematics: “What is the impact of participation in Math Buddies on student’s mathematics problem-solving achievement?” (p. 30).

Math Buddies was a computer based course that integrated mathematics, science and technology. Students used computers to problem solve. Nine teachers worked with university faculty during a summer to develop the course. Forty preservice teachers were trained in using the software and acted as tutors for the students. Data included a pre/post administration of an assessment co-developed by faculty and teachers that measured students’ ability to “understand, solve and extend the problem, and determine the kind of problem-solving strategy most appropriate” (p. 30). State test data were also examined. A t-test showed significant gains on the developed assessment for grades two ($t = 4.20$, $p = .001$), three ($t = 2.69$, $p = .01$), and four ($t = 2.38$, $p = .02$). Moreover, there was growth on the state mandated test for third and fourth grades, with third grade gaining seven percentage points and fourth grade gaining fifteen percentage points when compared to scores earned prior to participation in Math Buddies.

The authors acknowledge that this study is not generalizable, but they provide evidence that the study has: outcome validity, process validity, catalytic validity, and dialogic validity. It is somewhat lacking in democratic validity because it "omits the voices of parents and students" (Knight, Wiseman, & Cooner, 2000, p. 36). However, there were two serious limitations in this study; one is the treatment. Going to the lab for Math Buddies was an option for teachers; therefore, the students did not have equal exposure. Some teachers used the lab often, others very little. It was also available to students after school and students could choose to stay and work. Additionally, the assessment instrument was weakened after it was created. The team reworded the test using vocabulary similar to what the students would see on the state mandated test and changed the format of many of the items from open-ended to multiple choice. Therefore, the ability of the test to objectively measure problem-solving ability may have been compromised.

Leonard, Lovelace-Taylor, Sanford-DeShields, and Spearman (2004) investigated whether there was improvement in mathematics in two elementary PDSs that were part of a large urban university's PDS collaborative. This study was a reexamination of the PDS model that was adopted in 1991. The research questions were: "How are teachers teaching mathematics now? and, How can the PDS partnership effect improved mathematics instruction?" (p. 571). These questions were pursued because after a long history of collaboration, there was no improvement in mathematics for any of the schools that were partners. In order to answer the questions, case study research was utilized and university faculty members were participant observers in the PDS classrooms.

Faculty observed two veteran elementary teachers over the course of one month as they engaged their students in problem-solving. One teacher taught a class of primarily Latino students while the other teacher's students were African American. There were 26 students in each class. Data sources included audio and script tapes, and weekly observations. "The audiotapes and script tapes were analyzed to determine the teacher's style of instruction and the kinds of strategies students used to solve problems" (Leonard, Lovelace-Taylor, Sanford-DeShields, & Spearman, 2004 p. 571).

Findings showed that teaching styles varied between teachers. One teacher engaged in reformed teaching, encouraging students to try various methods, collaborate, and model problems for fellow-students. The other teacher's style was more in line with traditional teaching, with the teacher acting as the locus of control. Although both teachers had strong content knowledge, their teaching styles were very different. These findings are important because they illustrate that while mastery of content is crucial for effective teaching, it must be delivered through pedagogy that allows students to build their knowledge through inquiry, trial, and error.

Understanding the importance of the PDS model as a viable archetype to transform education, the National Council for Accreditation of Teacher Education (NCATE) developed a set of Standards "that would strengthen and support PDSs, as well as be used to assess their progress" (Teitel, 2004, p. 1). The Standards attend to the following key components of PDSs: "learning community…accountability and quality assurance…collaboration…diversity and equity…[and] structures, resources and roles" (p. 30). NCATE spent five years developing and field testing the Standards. They serve as a "framework for conducting and evaluating research that addresses the question of what outcomes can be attributed to involvement in PDS partnerships" (NCATE, 2001, p. 2).

Guided by these Standards, Teitel (2004) conducted a limited review of studies that examined the effectiveness of PDSs. He developed a framework for selection based on the following criteria:

> I. Is it a PDS? Do the authors use the PDS Standards as a framework and for self-assessment? II. If not, does the study tell us enough about the settings to do a rudimentary checklist of how the partnership might meet the PDS Standards, and where it falls on the developmental continuum? III. Are there a range of outcome measures that are logical and credible? IV. If the study is comparative, is there an appropriate balance between process and outcome and a theory of change in place—a logical suggestion that PDS approaches to teaching and learning could plausibly be leading to the outcomes that are noted? (p. 8)

The studies focused on changes in: preservice teachers, practicing educators, and K-12 student outcomes. The majority of studies that qualified for the analysis examined effects on preservice and practicing educators; however; for the purpose of this review, only the studies examining student outcomes in mathematics will be discussed.

The first study was conducted by RAND (2000). It was a preliminary evaluation of the Benedum Collaborative Model, a comprehensive PDS partnership developed by West Virginia University's College of Human Resources and Education. The collaborative included 21 public schools across five counties. The Model has been in existence since 1988. Using existing data from the Sanford 9 (SAT 9) test, the evaluation team investigated the following questions with respect to student achievement: "1. Do PDS students have higher test scores? 2. If PDS students score better, is it because their schools are better? 3. If PDS schools are better schools, is it because they are PDSs?" (Gill & Hove, 2000, p. 13).

When the PDS students' gain scores on the SAT 9 in basic skills in mathematics were compared to non-PDS schools in the five county area, it was found that the PDS students' gain scores were significantly higher, by 0.11 standard deviations, than non-PDS student scores at all grade levels. In order to determine whether there was any correlation between membership in the PDS and higher gain scores, the evaluation team conducted a natural experiment. "A natural experiment is a situation in which two groups begin in comparable circumstances and are then given different treatments" (Gill & Hove, 2000, p. 19).

The authors compared the gain scores of students who attended a non-PDS middle school followed by a PDS high school to students who attended both non-PDS middle and high schools. The groups consisted of students who had similar scores on the SAT 9 in middle school. Students' scores were compared after completing the second year of high school. It was found that the students who attended a PDS middle school had an average gain score of 17.5 scaled points compared to 10.5 scaled points for the non-PDS schools. Although not significant, this amounts to approximately "two-tenths of a standard deviation" (Gill & Hove, 2000, p. 20).

There are limitations to these results. No baseline data were collected for the first analysis; therefore, it is not possible to say whether the PDS students were already achieving at a higher level than non-PDS students prior to joining the collaborative. Additionally, no attempt was made to match schools in the five-county area with the PDSs. However, in comparing the gain scores between newer PDSs and older PDSs, it was found that the longer the school had been a PDS, the larger the gain scores were over time; however these gains were not significant. This is important because it illustrates that improvement in student achievement takes time and the sustainability of a PDS beyond a few years may be crucial to effect significant positive change.

In the second study, Pine (2003) examined changes in student achievement on the Michigan Educational Assessment Program (MEAP) test for the Longfellow Community Elementary School, a PDS associated with Oakland University in Michigan. The student population included: 70% African American, 25% White, and the remaining 5% a combination of Asian, Native American, and Latino/a. The Longfellow-Oakland partnership formed during the spring of 1991. During its first year as a PDS, 25.6% of fourth graders scored at least satisfactory on the mathematics portion of the MEAP. This was lower than the average score for the entire state of Michigan by several percentage points but slightly higher than the average score for schools in Pontiac, where Longfellow is located. By 1995, the average score for the PDS rose to 92.1% and to 100% by 2000. Scores had been consistently higher than both comparison groups throughout this time period.

Although it appears that involvement in the PDS model contributed to student gains, Pine conducted a secondary analysis on the artifacts generated during the time of the partnership. These included: “planning documents, school improvement plans, mission statement, newsletters, annual reports, correspondence, minutes of various meetings, and action research studies” (Pine, 2003, p. 44). One common theme emerged in the documents: “an overriding concern for student learning” (p. 44). This study is important because it suggests that a sustained focus on improving student outcomes, examining student work, and modifying instructional practices may have a positive effect on outcomes, especially for African American students.

Castle, Arends, and Rockwood (2008) studied the impact of a PDS on student achievement, comparing achievement levels of PDS students to a control school. Schools were matched on achievement levels and demographic variables. The research questions guiding the study were: “What is the impact of the PDS model on student learning? Specifically, is there a difference in student learning over time between a PDS and a control school?” (n. p.).

The PDS was a high needs school with: 80% minority students, 15% ELL students, and 26% transient students. Fifty-nine percent of the students received either free or reduced lunch. Less than 50% of 4th-grade students scored at the satisfactory level on the reading portion of the state mandated test. The control school was matched on all of the variables with a difference of no more than 4% on any measure. As well, comparisons were made between the PDS and all students in the district where it was located.

Data sources included scores for four cohorts of students during fourth and sixth grades on the mathematics portion of Connecticut Mastery Test (CMT), the state mandated test, over a period of six years. Results showed that scores declined from fourth to sixth grade for every cohort from each group: PDS, control, and district, except for PDS cohort 2, which experienced an increase of 4%. However, the level of decline was much lower for the PDS school than the other groups except for cohort 1, which had a decline of 18%. Overall, the mean score for all PDS cohorts dropped by 9.75%. This was less than the drop for the control school (15%) as well as the mean drop for all district schools (17.75%).

A finer grained analysis was completed to determine whether there was any improvement for students who were most at risk for failing. The PDS reduced the number of students receiving intervention by 10% compared to the whole district, which had a reduction of 1% and the control school, which increased the percentage of students receiving intervention by 1%.

Although the results show that the PDS did better than either the control school or the district as a whole, there are limitations to this study. The only comparisons made were in measuring gain scores for students from fourth to sixth grade. Furthermore, no statistical analyses were completed for the study. These results suggest that the PDS model may benefit students who are at risk for failing; therefore, a more comprehensive analysis, which tracks students over a period of time longer than two years, is warranted to determine whether the student achievement in the PDS continued to improve.

Houston, Hollis, Clay, Ligons and Roff's (1999) research on the Houston Consortium constitutes the final study reviewed by Teitel that will be discussed. Partners in the Consortium include: three school districts, two intermediate units, and four Institutes of Higher Education. Together the partners in the Consortium developed a program specifically targeting all of the prospective teachers attending the four partner universities. These teachers were being trained to teach in urban schools in Houston. Houston Independent School District, the fourth largest school district in the nation at the time the study was completed, was responsible for educating 30% of all African American students in the state of Texas. The Consortium articulated six primary goals for its PDSs, one of which was "to increase student learning in urban PDSs" (p. 9).

The outcome measurement that assessed improvements in student achievement for this study was the Texas Assessment of Academic Skills (TASS). "The study compared achievement for mathematics, reading, and writing over a 3-year period" beginning with 1992 (Houston et al., 1999, p. 25). Baseline scores for 1992-93, the year prior to the schools becoming PDSs, were compared to scores for 1994-95. Student outcomes included both elementary and secondary scores. Scores increased in mathematics for all 16 PDS schools compared to increases in reading for 14 PDS schools and increases in writing for 10 PDS schools.

The authors make no claim for causality but point to unique features of the Consortium that may have helped to contribute to student gains:

> - A larger number of adults were involved in instruction with the addition of preinterns and interns in PDSs. The student-adult ratio was thus lowered, providing greater access by each student to adult mentoring, support, tutoring, and instruction.
> - Teachers had opportunities to engage in professional development related to teaching in urban schools and the use of technology.
> - School leaders were engaged in a network in which personnel from schools, intermediate educational agencies, and universities had opportunities to solve problems and share ideas.
> - Being part of a new movement selection after competitive applications to be PDSs, and the Hawthorne effect of change all were factors that may have influenced achievement. (Houston et al., 1999, p. 26)

The PDS model seems to have sustained the test of time; this is indicative of the continual commitment of partnerships between K-12 schools, universities and the teachers they prepare to improve education for the nation's students. The body of research on the PDS model will most likely continue to grow; using the NCATE Standards as an evaluative tool will increase the rigor of those studies.

Professional Development: A Mainstay for School Improvement

Professional development of teachers is not unique to large scale reform efforts or PDSs. Nearly every state in the union mandates that teachers continue to grow professionally. Whether large scale or school based, professional development has been the focus of much research over the last several years. Guskey (2002) defines professional development in the following way: "…programs which are systematic efforts to bring about change in classroom practices of teachers, in their attitudes and beliefs and in learning outcomes of children" (p. 381). In 2001, the Committee on Science and Mathematics Teacher Preparation published a seminal report which found that:

> After extensive review of the research literature and the recommendations of professional societies, the National Research Council's Committee on Science and Mathematics Teacher Preparation (CSMTP) has determined that fundamental restructuring of teacher preparation and professional development is needed to best serve the interest of students' learning and of their future success as individuals, workers, and citizens. (CSMTP, 2001, p. 2)

The Committee called for "using evidence and analysis to build an effective system of teacher preparation and professional development" (p. x). Their recommendation has two implications. One is careful scrutiny of teacher preparation and professional development, and the second is ongoing observation and evaluation of teachers' instructional and assessment practices.

In January of 2006, the U.S. Department of Education issued a report, *Teacher Professional Development in 1999-2000*. This report highlights the findings garnered from the Schools and Staffing Survey (SASS). Respondents to this survey included "12,300 public and private school principals and 52,000 public and private school teachers" teaching in 12,000 schools in 4,700 districts nationwide (Choy, Chen, & Bugarin, 2006, p. 4). The report reveals information on two levels of professional development. "The first part of the report examines the ways in which districts and schools organize and manage professional development, and the second part describes the extent to which teachers participate in various types of professional development activities" (p. 1). Several findings emerged. Professional development decisions typically occur at the department, school, or district level and are largely determined by what is perceived as teacher and/or student need; however, as the size of the staff increased, the amount of influence teachers had in professional development decisions decreased. Professional development activities are usually conducted by outsiders and nearly 90% of principals stated that professional development was closely tied to academic standards and/or district/school improvement.

In response to the increasing demands of NCLB, policymakers have identified professional development as the locus of reform initiatives (Corcoran, Shields & Zucker, 1998). "Five streams of reform, both singularly and in combination, present complex challenges to teachers as individuals" as they engage in professional growth (Little, 1993, p. 130). They are:

> reforms in subject matter (standards, curriculum and pedagogy), reforms centered on problems of equity among a diverse student population, reforms in the nature, extent, and uses of student assessment, reforms in the social organization of schooling, and reforms in the professionalization of teaching. (p. 129)

Professional development activities often have a particular focus which may relate to one or more reform initiative. In response to the literature of late, professional developers have designed activities that help teachers increase content knowledge. Some activities focus on improving pedagogy, while others support teachers as they implement new curricula, or the focus may be on student understanding. Often, what begins as a movement toward reform ends up being a cafeteria style or workshop oriented professional development program. These have often proven ineffective in instituting change (Garet, Porter, Desimone, Birman, & Yoon, 2001; Little, 1993; Wilson, 2003).

Garet et al. (2001) identified some core features of professional development that have positively impacted teachers in increasing pedagogical content knowledge and transforming classroom practice. Utilizing an ordinary least squares regression model on "a national probability sample of 1,027 mathematics and science teachers", teachers self-reported the following as contributing to their learning: "(a) focus on content knowledge; (b) opportunities for active learning and; (c) coherence with other learning activities" (pp. 915-916). Structural features also emerged that contributed to teacher growth: "(a) the form of the activity (e.g. workshop vs. study group); (b) collective participation of teachers from the same school, grade or subject; and (c) the duration of the activity" (p. 916). The authors then examined the relationship between the changes in teacher knowledge and practice and concluded:

> Activity type has an important influence in duration; reform activities tend to span longer periods and to involve greater numbers of contact hours than traditional activities. Our results also show a modest direct effect of activity on enhanced knowledge and skills, indicating that reform activities have slightly more positive outcomes when all of the design features and quality characteristics in our model are included. Our two measures of duration—time span and contact hours—exert a substantial influence on the core features of professional development experiences. Time span and contact hours have a substantial positive influence on opportunities for active learning and coherence. (Garet et al., 2001, p. 930)

Whatever the focus, "professional development activities frequently are designed to initiate change in teachers' attitudes, beliefs and perceptions" (Guskey, 2002, p. 382). These changes will lead to changes in practice, which will help to improve student learning. However, this trajectory oversimplifies what is involved in changing attitudes and beliefs and it assumes that a programmatic innovation, when delivered by an experienced professional, has the power to influence attitudes and beliefs so that change will occur. Guskey (2002) developed an alternative model of professional development which “portrays a temporal sequence of events from professional exercises to enduring change in teachers’ attitudes and perceptions” (p. 382). This model places changes in teachers’ beliefs and attitudes at the end of trajectory, suggesting that these changes occur when and only when teachers see positive changes within their classrooms, either in their teaching or in student outcomes that are directly related to the professional development that they received.

Similar to the research on PDSs, the research on how professional development affects student achievement is sparse. Four hundred fifty professional development projects were reviewed by the Middle Grades Initiative program. Of these, less than 10% could provide evidence of having any impact on student performance. “Most of these projects simply weren’t designed to collect evidence, or lacked sufficient funds and resources to do so” (Killion, 1998 n. p.). Twenty-five of the projects were investigated further to see how closely the projects adhered to four norms. Briefly stated, they are:

> (1) The program measures results based on student performance; (2) The staff development is well organized; (3) The program offers content specific development for middle school teachers; (4) The program is being used at multiple schools, across districts, statewide or across a regional area. (n. p.)

Additionally, each of the 450 programs was evaluated on four levels: participants’ reaction to and satisfaction of the program; “changes in participants’ attitudes, content knowledge, or knowledge of skills”; changes in practice and level of implementation;

and, “the impact of teacher learning on student performance” (n. p.). Very few of the programs provided evidence of increased student performance.

Linking professional development activities to increases in student achievement has continued to prove elusive, perhaps due in part to the many factors that can influence achievement. These include but are not limited to: teacher/student ratio, available resources, years of teaching experience, and existing support systems for struggling students (Killion, 1998). Or perhaps, as Guskey (1997) suggests:

> There are three particularly notable reasons why past efforts to identify the elements of effective professional development have not yielded more definitive answers: (1) confusion about the criteria of effectiveness, (2) the misguided search for main effects, and (3) the neglect of quality issues. (n. p.)

Historically, the effectiveness of professional development programs has been based on self-reports of teachers who participated in them. “This information tends to be highly subjective and not particularly reliable” (n. p.). Other factors included the lack of clear documentation about the knowledge and skills teachers gain and how they actually implement this new knowledge in their classrooms, and the use of student data on learning outcomes. These can be informative in evaluating professional development, but they are often difficult to measure because there is a lack of specific criteria to measure against.

Yoon, Duncan, Lee, Scarloss, and Shapley (2007) reviewed more than 1,300 studies claiming to document changes in student achievement associated with professional development. Holding the studies to the What Works Clearinghouse Evidence Standards, the authors identified only nine studies that met the standards for methodological rigor. Five studies, which examined achievement in mathematics, are summarized here.

Ross, Bruce and Hogaboam-Gray (2006) investigated the following question: “Does teacher professional development enhance student achievement in mathematics?” (p. 557). The study was a randomized field trial conducted in one district. The population under investigation included all the teachers who taught sixth grade (n = 120). Teachers were randomly assigned to an early or late professional development program. The sessions occurred during September through December for the early group (treatment), and January through May for the late group (control). Six students from each class were chosen randomly as the participants. In order to test for the equivalency of groups, students completed a survey in September that measured the following constructs: “(i) mathematics self-efficacy…(ii) task goal orientation…(iii) ability-approach goal orientation…(iv) ability-avoid goal orientation…(v) negative affect for failure… (vi) classroom mastery goal structure…(vii) classroom performance and goal structure [and] (viii) effort” (pp. 560-561). A t-test revealed no significant difference among students.

The professional development occurred during one school day with three after-school gatherings lasting two hours over the course of 10 weeks. The overarching theme of the PD was communication focusing on the following: “(i) the use of rich tasks…(ii) sharing and appraising mathematical ideas in student groups and whole class discussion, and (iii) teachers and students collaboratively constructing mathematical knowledge” (Ross, Bruce, & Hogaboam-Gray, 2006, p. 561). Experimental and control groups received the same professional development, with the treatment occurring during September for the experimental group and during January for the control group.

Data sources included a pre-test/post-test assessment created by a team of teachers who authored the yearly mandated test. It assessed skills in number sense and numeration, and patterns and algebra. The assessment was field tested in two similar districts on a sample population of 140 students. The assessments "generated eight scores: four aspects of mathematics achievement (problem solving, concept understanding, application of mathematical procedures, and communication of mathematical ideas) × two strands of mathematics (Number Sense & Numeration and Patterning & Algebra)" (Ross, Bruce, & Hogaboam-Gray, 2006, p. 558). These were further collapsed into questions that measured either procedural knowledge or conceptual knowledge. The authors claimed that the instrument was validated through content analysis and confirmatory factor analysis, but they provide no psychometric data to support this claim.

In order to address reliability issues, all assessment scores were verified through inter-rater reliability. When discrepancies were not negotiated, a third party, who was a master scorer of the state mandated test, was consulted. In order to provide a second level of reliability, items were drawn randomly at the beginning, middle, and conclusion of the marking session and scored by an additional scorer (Kappa = .73, .97, and .97). The authors attended to validity in two ways. Internal validity was substantiated by the use of random assignment and the establishment of similar groups. External validity was substantiated by "including all grade 6 teachers in the district" (Ross, Bruce, & Hogaboam-Gray, 2006, p. 558).

The dependent variable was the score on post-test assessment. The independent variable was the treatment, and the covariate was the score on the pre-test. An ANCOVA revealed a "statistically significant pre-test effect [$F(1, 103) = .111$, $p < .001$]"; however, there were no differences between groups on the post-test (Ross, Bruce, & Hogaboam-Gray, 2006, p. 567). The authors also evaluated improvement over time, comparing the students' performance on the yearly mandated test between 2003 and 2004. In this analysis, control and treatment groups were combined because by this time all the teachers had received the professional development. The authors found achievement levels in mathematics increased significantly for the whole district ($t = 3.73$, $p < .001$). The annual test also included assessments in reading and writing. When these scores were examined for growth, there were no significant differences from prior scores.

There is little generalizability in this study because it involved only one school district; however, the authors make the following recommendations:

> Intensive experiences with selected teachers in ideal circumstances could constitute existence proofs for teacher change. Qualitative case studies could identify processes in these sites most plausibly linked to improved teaching. Large scale PD activities would involve short duration, less intensive learning opportunities based on findings from existence proof studies. Annual achievement mentoring would measure progress over time. (Ross, Bruce, & Hogaboam-Gray, 2006, p. 573)

Saxe, Gearhart, and Nasir (2001) conducted a study examining elementary students' procedural and conceptual knowledge of fractions. The purpose of the study was "to understand the ways that professional and curricular supports for reform implementation may strengthen students' developing knowledge of fractions" (p. 57).

Participants included 23 elementary teachers who were invited to be part of the study. Teachers were divided into two groups, traditional or reformed, based on their responses to application questions. Traditional teachers constituted one group; reformed teachers were then separated into two groups through stratified random assignment, The Integrated Mathematics Assessment (IMA) group or the Collegial Support (SUPP) group. The IMA group received professional development designed by the authors. Consulting the findings from existing research on teaching and learning, the training focused on four themes: "(a) teachers' understanding of the mathematics that they teach…(b) teachers' understanding of children's mathematics…(c) teachers' understanding of children's achievement motivations in mathematics…and (d) the opportunity for teachers to work with other professionals concerned with effective implementation of reform (Saxe, Gearhart, & Nasir, 2001, pp. 59-60). Nine teachers participated in a summer institute that lasted five days with follow-up meetings every two weeks during the school year.

In the Support Program (SUPP) group, the teachers' work centered on implementing the reformed fractions units. Eight teachers reflected on their teaching practices together. One full day was devoted to each unit with follow-up monthly meetings taking place in the evenings during the school year. Although some of the IMA themes surfaced during the SUPP group sessions, no training was received; the themes were simply topics of discussion as teachers reflected on their practices. The Traditional (TRAD) group (n = 6) received no treatment; their only requirement was that they be committed to a traditional method of teaching.

Students in grades four and five completed a pre-test prior to beginning a unit on fractions and a post-test upon its completion. Student demographics consisted of 64% Latino, 14% White, 8% African American, 7% Asian and 7% Other. English fluency was unevenly distributed across groups; therefore, the authors "adjusted for language background statistically in [their] analyses" (Saxe, Gearhart, & Nasir, 2001, p. 66).

The assessment instrument was created by the authors and contained items that measured knowledge in fraction computation as well as conceptual knowledge. The items were drawn from materials that were currently in use in the locale where the study was completed. These items were pilot tested for level of difficulty and ambiguity prior to administration. Three subscales were created: general fraction knowledge, computational knowledge, and conceptual knowledge. Cronbach's alpha revealed indices of 0.73 for computation and 0.86 for conceptual knowledge for the pre-test, and 0.83 and 0.8, respectively, for the post-test, suggesting relatively high internal consistency for both instruments. Confirmatory factor analyses indicated that the subscales for computation and conceptual knowledge were indeed independent measures. For the pre-test, the chi-square value ($df = 11$) was 17.254, $p = 0.10058$, and ($df = 11$) 19.1, $p = 0.059$ for the post-test.

> Analyses of the pre-test data indicated a linear dependency, a problem that probably resulted from a heavily skewed distribution towards the floor, given lack of prior instruction in fractions. Overall, our analyses indicated that the two sets of items were indexing independent areas of competence in children. (Saxe, Gearhart, & Nasir, 2001, p. 67)

The authors found that students in all groups (IMA, SUPP, and TRAD) showed growth from the pre-test scores. Further analyses were completed to determine whether there were differences between groups on the post-test. Results for the ANCOVA for the conceptual scale showed:

> a main effect for GROUP ($F(2,18) = 7.21$, $p < 0.005$). The overall means and standard deviations of groups for IMA, SUPP, and TRAD were 6.17 (0.89), 4.73 (1.0), and 4.10 (0.68), respectively. Tukey-HSD post hocs

> ($p < 0.05$) on adjusted scores revealed that the IMA classroom means were greater than both the SUPP and the TRAD classroom means. (Saxe, Gearhart, & Nasir, 2001, p. 69)

These findings suggest that the use of a reformed curriculum, when supported through focused professional development may lead to gains in student achievement that are superior to gains for students whose teachers are either traditional or teachers who use reformed curricula without the support of professional development. The current study is informed by these findings in that it investigates increases in student outcomes for students in high schools where both traditional and reformed curricula were used and where teachers' participation in professional development differed in intensity.

Carpenter, Fennema, Peterson, Chiang, and Loef (1989) investigated whether student achievement was enhanced when teachers were exposed to research on how children understand addition and subtraction. The following research questions were pursued:

> 1. Did teachers who had participated in a program designed to help them understand children's thinking (a) employ different instructional processes in their classrooms than did teachers who had not participated in the program? (b) have different beliefs about teaching mathematics, about how students learn, and about the role of the teacher in facilitating that learning than did teachers who did not participate in the program? and (c) know more about their students' abilities than did teachers who did not participate in the program? 2. Did the students of teachers who participated in a program designed to help them understand children's thinking (a) have higher levels of achievement than did the students of teachers who did not participate in the program? (b) have higher levels of confidence in their ability in mathematics than did the students of teachers who did not participate in the program? and (c) have different beliefs about themselves and mathematics than did students of teachers who did not participate in the program? (p. 503)

Participants included 40 first-grade teachers who were randomly assigned to either a treatment group (n = 20) or an experimental group (n = 20). The experimental group received training in Cognitively Guided Instruction (CGI) over four weeks during the summer. The intent of the professional development was to expose teachers to research on young children's problem-solving abilities in addition and subtraction. The control group took part in two workshops that focused on non-routine problem-solving; however, there was no dialogue or training in how students think or learn. Each workshop lasted two hours.

Data sources for students included scores on the Iowa Test of Basic Skills (ITBS) (Level 6), administered during September and follow-up post-tests administered during April and May. Post-tests included subscales of the ITBS (Level 7), which measured computation and problem-solving ability. Reliability was high on all measures. Cronbach's alpha ranged from .72 to .91. Additionally, interviews were conducted with students as they engaged in problem-solving. Students also completed surveys, which measured attitudes and beliefs. These were developed specifically for the study. Six boys and six girls were randomly chosen from each classroom to serve as the sample to be observed and interviewed.

Data sources for teachers included: classroom observations, interviews, and questionnaires completed the academic year following the summer treatment. Teachers' knowledge of their students was measured near the end of the school year. During the interviews teachers were asked to make conjectures about how students would solve problems and whether their methods would lead to correct answers. "Teachers' predictions were then matched with students' actual responses to obtain a measure of teachers' knowledge of their students' thinking and performance" (Carpenter et al., 1989,

p. 503). Teachers also completed a survey which measured their "assumptions about learning and teaching" (p. 503).

Two types of observations occurred; one focused on the teacher while the other focused on the target group of students. Teachers and students were observed over the course of one week, four times between November and April. The observation protocol for teachers contained six categories: "setting, content, expected strategy, teacher behavior, process focus, and answer focus". The protocol for students included: "setting, content, strategy used, and lesson phase" (Carpenter et al., 1989, p. 509). Inter-observer agreement ranged from 82% to 99% for both teacher and student observations.

A 12-item questionnaire was completed by the teachers that assessed their beliefs about how students learn addition and subtraction. The instrument contained four scales:

> The Role of the Learner…Relationship Between Skills…Sequencing of Mathematics…[and] Role of the teacher….Cronbach alpha estimates for teachers' scores in this sample (N = 40) were .93 for the combined scales and .81, .79, .79, and .84 for Scales 1 through 4, respectively. (Carpenter et al., 1989, p. 512)

The authors analyzed the following data: student achievement, teacher knowledge and beliefs, and classroom observations. The unit of analysis was either the teacher or the classroom. Results for classroom observations showed a significant difference between CGI and control teachers in: number fact problems, $t(38) = -3.77$, $p < .05$; word problems, $t(38) = 2.84$, $p < .05$; teacher poses problem, $t(38) = 3.76$, $p < .05$; teacher listens to process, $t(38) = 3.54$, $p < .05$; teacher gives feedback to process, $t(38) = 3.54$, $p < .05$; teacher gives feedback to answer, $t(38) = -2.45$, $p < .01$; advanced counting as an expected strategy, $t(38) = -2.04$, $p < .01$; and, the use of multiple strategies, $t(38) = 2.45$, $p < .01$. Results for the test of teachers' knowledge showed significant differences in teachers' knowledge about their students use of number facts, $t(38) = 3.51$, $p < .01$, and problem-solving abilities, $t(38) = 3.56$, $p < .01$. ANOVA results for teachers' beliefs showed that "CGI teachers were significantly more cognitively guided in their beliefs than were control teachers" (Carpenter et al., 1989, p. 523). In the area of student achievement, results showed no difference on the ITBS (Level 7) in computation; however, a significant difference did emerge in the recall of number facts for CGI students, $t(37) = 2.23$, $p < .05$. There was also a significant difference in complex addition/subtraction ability, $t(37) = 1.99$, $p <. 05$, one-tailed test.

This study is important because it demonstrates that professional development that is focused on research about how students learn can significantly impact teacher's practices and students' use of problem-solving strategies. It also suggests that reformed teaching can happen with traditional curricula. It informs the present study in that some of the partner districts in the MSPGP continued to use traditional curricula, but many participated in professional development in reformed teaching practices.

Two follow-up studies were conducted that support the findings of Carpenter, Fennema, Peterson, Chiang, and Loef. During 1993, Villaseñor and Kepner investigated whether professional development grounded in the CGI philosophy benefited students in urban classrooms. The authors conducted a quasi-experimental study in 24 first-grade classrooms where at least 50% of the students were minorities. Twelve classrooms were randomly selected as the experimental group, while the remaining 12 classrooms served as the control group. Minority populations ranged from 57% to over 98% in the experimental classrooms and from 58% to over 99% in the control classrooms.

The treatment consisted of a 19-hour CGI workshop completed during the summer. Teachers attended follow-up sessions for two hours during the school year, one occurring during October and the other during December. The professional development consisted of problem-solving activities but did not include “CGI principles or research” (Villaseñor & Kepner, 1993, p. 65). First-grade students completed a pre-test during October and post-tests during February and March. A random sample of six girls and six boys from each of the 24 classrooms was chosen for the analysis. The test included number fact questions and word problems. Upon completion, students were interviewed about their problem-solving strategies and knowledge of number facts. Prior to the student interviews, a total of 82 observations were completed in both the treatment and control classrooms using the observation guide developed by Carpenter, Fennema, Peterson, Chiang, and Loef (1989). These observations informed the development of the interview questions used in the study.

Results showed a significant difference between the CGI and control students on the pre-test $t(22) = 2.98$, $p < .01$. After adjusting for this difference, it was found that the CGI students scored better than the control group on the post-tests. Using the student interviews, a finer grained analysis was done in order to determine whether CGI students used higher-order strategies to solve both the number fact and word problems. Again the CGI students outperformed control students. Out of a maximum score of 6 for problem-solving, CGI students’ mean score was significantly higher (3.00) when compared to the control group (.67), $t(22) = 4.29$, $p < .01$. When number fact problems were compared, it was found that out of a maximum score of 5, the mean score for CGI students was again significantly higher (3.65) when compared to the control group (.90), $t(22) = 5.45$,

$p < .01$. This study supports the work of Carpenter, Fennema, Peterson, Chiang, and Loef (1989) and suggests that problem-solving skills can be developed in children even when lower order skills have not yet been fully developed.

During 1996, Fennema, Carpenter, Franke, Levi, Jacobs, and Empson completed a four-year longitudinal study which was a follow-up study to the one completed during 1989. However, in this investigation the student population included students in first, second and third grades. The authors explored the following questions:

> What knowledge will enable teachers to modify their instruction so that it becomes more inline with current recommendations? A related and perhaps more important issue is what students learn when instruction changes as a result of teachers' new knowledge. The study reported here addresses these questions. (pp. 403-404)

Twenty-one teachers participated in the study focused on how teachers' instructional practices and beliefs changed with respect to increased knowledge about how students think. Baseline data were collected on both the teachers and the students during Year 0. The professional development that occurred during the next three years included CGI training and support for teachers in their classrooms as they applied this new knowledge to their teaching. Data collection methods included: observations, post observation interviews, The CGI Belief Scale, and "informal interactions" (Fennema et al., 1996, p. 405).

The sample consisted of 21 of the 65 teachers who participated in the professional development. All of the teachers were female, teaching at three different schools that had a predominately white student population. The percentages of students eligible for free or reduced lunch were: 4%, 26%, and 8%.

During the baseline year, a workshop was provided for teachers who had little or no knowledge of CGI. It lasted 2.5 days. The following year, all teachers attended a 2-day workshop during the summer. This was followed by 14 workshops during the academic year, each lasting three hours. Year 2 treatment included two workshops lasting half an hour and a 2-day workshop focused on teacher reflection. During the final year, teachers participated in one 3-hour reflection workshop and in two half-hour review workshops. Ongoing support consisted of classroom visits by trained CGI staff. Weekly visits occurred during Year 1. The visits decreased to every two weeks during Year 2 and only occasionally during Year 3.

In order to assess Teachers' CGI beliefs and instructional practices, rubrics were developed that identified various levels of beliefs and instruction. Data sets for teachers included: "transcriptions of all interviews, observations and field notes" (Fennema et al., 1996, p. 413). All data were collected prior to any analysis. These data were analyzed by five staff members who helped to develop the rubrics. Children's learning was assessed with a test that measured both conceptual and problem-solving knowledge:

> Estimates of internal consistency of the Concepts and Problem-Solving tests, estimated using Cronbach's alpha, were .84, .81, and .77 for the first-, second-, and third-grade tests, respectively. For both the second and third grade, Computational Skills tests were also developed with Cronbach's alphas of .71 and .73, respectively. (p. 413)

The unit of analysis for measuring changes in student knowledge was the mean score for each teacher's classroom. It was not possible to calculate gain scores for students because teachers taught different students each year.

Results showed that 19 of the 21 teachers' instructional practices changed to reflect CGI principles during the three years. This finding was supported by the results of the CGI Belief Scale. “A one-tailed t-test on teachers' scores on the CGI Belief Scale showed a significant improvement from the initial year (M = 151.19, SD = 22.63) to the final year (M = 164.38, SD = 18.16, $t = 3.74$, $p < .01$)” (Fennema et al., 1996, p. 421). However, no clear pattern emerged when the rubrics measuring levels of instruction and beliefs were compared. When students’ test scores were examined, it was found that mean scores improved in every teacher’s classroom.

Overall, these results suggest that a change in a teacher's level of instruction was reflected in the achievement of her students. “Changes in a teacher's level of instruction tended to have the greatest influence on student achievement the year following the change” (Fennema et al., p. 429). This study is important because it illustrates that as teachers moved toward CGI instructional practices, they were more focused on problem-solving than on computation. However, even though less emphasis was placed on computation, students’ scores did not suffer, and in fact, computation mastery increased. This informs the current study because one of the unique features of the MSPGP was implementing reformed curricula that focused on student thinking. Although these curricula are not CGI, the philosophy is similar in that each utilizes student’s thinking to develop mathematical literacy.

The Council of Chief State School Officers (CCSSO), with funding from NSF, completed a meta analysis during 2009 on studies that claimed to show evidence that there is a relationship between professional development and student achievement. The questions guiding the analysis were:

> 1) What are the effects of content-focused professional development for math and science teachers on improving student achievement as demonstrated across a range of studies?

> 2) What characteristics of professional development programs (e.g., content focus, duration, coherence, active learning, and collective participation of teachers) explain the degree of effectiveness, and are the findings consistent with prior research on effective professional development (e.g., content focus, duration, coherence, active learning, and collective participation of teachers)? (Blank & de las Alas, 2009, p. 5)

The authors employed a logic model to systematically accept or reject studies for their analysis. Of the 416 studies that were reviewed, 74 were initially accepted for further examination. Each was coded and analyzed by a team of doctoral students. A study was accepted if it:

> 1) presents an empirical study with quantitative data on an in-service professional development program for teachers of math and/or science and includes student achievement outcomes; 2) uses a research design that produces valid and measurable results; 3) reports at least one effect size or provides sufficient data to compute at least one effect size; and 4) records some professional development characteristics. (p. 8)

Inter-rater reliabilities were: 0.81 for round one; 0.91 for round two; and, 0.95 for round three. In the final round, 16 studies were accepted. Twelve studies investigated mathematics. Five were randomized control trials (RCTs); seven were quasi-experimental design studies (QEDs). Three of the 12 focused on high school achievement.

The analysis revealed that RCTs produced a greater mean effect size (M = .27) than QEDs (M = .17) for pre-test/ post-test studies. Different student measures also yielded different effect sizes. For studies using assessments aligned to the professional development program, the mean effect size was .32. For those that used state tests it was .01.

> For the studies that used a post-analysis only four types of achievement tests were found. The mean effect size for…program-specific student assessment[s] was .28,…[for] national norm-referenced assessments [it] was .17,…[for] local achievement tests [it] was .05,….[and for] statewide

> criterion referenced assessments [there was] a small mean negative effect size (-.07). (Blank & de las Alas, 2009, p. 22)

When characteristics of professional development were compared to changes in student achievement, a negative mean effect size (-.19) was found for professional development that focused on mentoring teachers. However, for programs that provided internships, the mean effect size was .20, and for those that offered "collaborative networking" the mean effect size was close to zero, suggesting that these types of professional development have little influence on increasing student achievement.

Increasing student achievement will continue to be at the forefront in PreK-12 education; NCLB has compelled school districts and teachers to improve student learning and provide evidence of that improvement. Furthermore, professional development will most likely continue to be the means to do this. However, linking professional development activities with changes in student achievement has been a challenge for educators and researchers. Blank and de las Alas' meta analysis is important because it systematically identifies rigorous research on successful professional development. This may be important for designing professional development programs in the future that will be effective. It also informs the education and research community about how to design research that produces valid conclusions on how to link professional development to changes in student outcomes.

Thus far, this review has looked at student achievement as an outcome of intervention. It will now focus on student achievement as an issue of equity. Inequitable education has perpetuated chronic underachievement for at-risk and minority students;

however, some progress has been made and this review will now discuss successful interventions that have occurred.

Student Achievement: The Perennial Problem

During 2009, the U. S. Department of Education released data on student dropout rates for the academic year 2006-07. On average, 73.9% of the entering freshman class of 2003-04 graduated on time. When this is disaggregated, 61.3% of Native and Alaskan Americans, 60.3% of African American and 62.3% of Latino/a students graduated on time compared to 91.4% of Asian students and 80.3% of white students (Stillwell, 2009). Although one of the goals of NCLB is dropout prevention, the overall graduation rate for 2006-07 was identical to 2002-03 with a slight increase during 2003-04, then dropping off during subsequent years.

When NAEP scores in mathematics are examined for 8th-grade students, achievement levels have continued to improve since 1978. There was a 14% increase for African American students, and for white students, a 6% increase. Although growth for African American students outpaced that of Whites, there was still a gap of 31 points during 2007 (Vanneman, Hamilton Anderson, & Rahman, 2009). In terms of proficiency, 13% of African American and 19% of Latino/a 8th-graders were proficient, compared to 55% of their white peers (NCES, 2010).

On the other hand, there is some good news. The "total minority enrollment at the nation's colleges and universities rose by 50 percent…between 1995 and 2005" (Cook & Córdova, 2007, p. 3). Moreover, the number of master's degrees earned by African Americans is nearly twice what it was during 1995, while the number of doctoral degrees conferred, grew by approximately 84%. For Latinos/as, doctoral degrees grew by 83% while master's degrees doubled. During 2005 "Students of color made up 19 percent of the nearly 17.5 million students on America's campuses" (Cook & Córdova, 2007, p. 2), compared to 13% during 1970 (NCES, 2009). Although this progress is a step in the right direction, it is slow and there is still much to be done. Only 7.1% of all doctoral degrees conferred during 2004 were earned by African Americans, and 41% of those were in education, "more than twice the percentage for whites" (Doctoral Degree Awards to African Americans, Winter, 2005/2006, p. 8). In the area of STEM education in the United States, only 16% of all bachelor's degrees earned were in STEM subjects compared to 64% for Japan and 52.1% for China, possibly leaving the U.S underprepared to compete globally in the future (National Science Board, 2006). For Minority students, the picture is even bleaker; only 10.6% of all minority students earned a degree in a STEM field during 2003. Moreover, less than 1% of all doctoral degrees in mathematics were earned by African Americans (Doctoral Degree Awards to African Americans, Winter, 2005/2006).

In order to turn the tide, NLCB has mandated that all students be 100% proficient in mathematics and reading by 2014. This assumes improvement overall, and specifically: "To close the achievement gap with accountability, flexibility, and choice, so that no child is left behind" (PL 107-110, 2002, p. 1). This is not easily done. "Raising student achievement generally and reducing achievement gaps are not the same thing….If equity is the goal, focusing on reform strategies that power higher achievement for all students will not ameliorate the gap" (Murphy, 2009, p. 11).

Narrowing achievement gaps has been the educational mission of schools, districts, and states for many years, as well as a major field of study for educational researchers. A keyword search for "achievement gaps" on EBSCOhost yielded 27,694 titles. Gloria Ladson-Billings (2006) offers a different lens through which to view achievement gaps. She suggests that over the long term, focusing solely on the achievement gap will not lead to sustainable improvement or long-term growth.

> Our focus on the achievement gap is akin to a focus on the budget deficit, but what is actually happening to African American and Latina/o students is really more like the national debt. We don't have an achievement gap; we have an education debt. (p. 5)

It is a debt rooted in history, economics, socio-politics, morality, and white privilege.

Gutiérrez (2008) classifies current research practices as "gap-gazing". Although contemporary research has moved away from simply identifying the existence of gaps to pointing to variables that may correlate with them, such as "income, family background….those variables are often not reasonable levers for change in the mathematics education community" (p. 358). Rather, she suggests that research remove the ceiling of white achievement as the goal to closing gaps and replace it with a focus on excellence in education. This way, the agent for change is how to increase and improve… "access,…achievement,…identity,…and power" for minority students to achieve excellence, making equity the focus. The focus on differences between groups fosters a mindset that "marginalized students are not worth studying in their own right—that a comparison group is necessary" (p. 359). In order to improve achievement of marginalized students, it is important to understand the reasons for success when it does occur. This includes the educational environment, resources, curricula, effective teachers, and the like.

Lubienski (2008) agrees with Gutiérrez in that simply identifying those factors such as socio-economics, etc. that appear to correlate with achievement is not enough, but she argues that “skillful analyses of gaps in students’ mathematics opportunities and outcomes can refute deficit perspectives by showing that student achievement is not fixed and by highlighting disparities in educational experiences and resources that contribute to achievement gaps” (pp. 352-353).

The remainder of this section will examine research focused on achievement and achievement gaps but will include only studies which examine relationships between variables that have the potential for change or that appear to influence achievement such as: excellent teachers, challenging courses, and access to resources and technology. Research suggests that these factors as well as opportunity to learn play a significant role in learning mathematics for underrepresented groups (Darling-Hammond & Post, 2000; Flores, 2007; Struchens, 2000).

Access and opportunity to study challenging mathematics is one road to success. Wang and Goldschmidt (2003) investigated the relationship between high school achievement and mathematics courses taken during middle school. The research questions considered were:

> 1. Is the mathematics course-taking distribution in Grade 8 different by students' gender, ethnicity, immigrant and language proficiency status? And how is this distribution related to the mathematics course-taking distribution in Grade 11? 2. Do middle school course-taking patterns have any independently long-lasting effects on students beyond middle school achievement? If so, what are the effects on students' high school (Grades 9-11) mathematics achievement? (pp. 3-4)

In order to answer the questions Wang and Goldschmidt drew a sample from a school district in California with a large heterogeneous population. The population included 3,346 students, 34% of whom were immigrants with various degrees of English proficiency. The sample included 2,707 students attending 17 different schools. Students were tracked over the course of four years. Scaled scores on the California Test of Basic Skills (CTBS) and the number and type of mathematics courses completed during high school were the student measures used in the evaluation.

Using Hierarchical Linear Modeling (HLM) to analyze the data, eighth graders taking remedial mathematics scored on average 26 points lower on the CTBS than students taking grade level courses. Students in advanced mathematics courses scored on average eight points higher than regular mathematics students. Further analysis showed that students in remedial classes who were also English language learners scored on average 14 points lower than students in remedial classes whose primary language was English. These results are troubling because they appear to have a direct effect on success in high school. When comparisons were made for subgroups of students, African American and Latino/a students were overrepresented in remedial courses. “By 11th grade, the percentage of Latino/a students enrolled in advanced mathematics courses was only about one quarter to one half of what the percentage should have been…for African American students [it] was far worse. There were only .14 as many African American students enrolled in advanced mathematics courses as there should have been” (Wang & Goldschmidt, 2003, p. 14).

This study illustrates that exposure to at least grade level mathematics has a positive impact during high school. Conversely, students who were assigned remedial mathematics during middle school did not catch up; this appears to directly contribute to the underrepresentation of African American and Latino/a students who enroll in challenging mathematics courses as well as those who enroll in grade level mathematics courses during high school.

McClain and Berry (2009) investigated the role that pedagogical style has on student achievement. The study included two middle school mathematics teachers. Fifty percent of the students were African Americans and 56% were eligible for free or reduced lunch. One teacher's pedagogical style was akin to the "initiation-response-evaluation (IRE)" style of teaching (p. 462), while the other teacher engaged in culturally relevant pedagogy. She had high expectations for her students and utilized the "lived and cultural experiences" of her students in her lessons (p. 461). Both teachers had over 10 years of teaching experience.

One hundred students in two groups were the participants. Pre-tests and post-tests were administered to all students. Each test consisted of eight multiple choice questions and two open-ended problems, which were drawn from the state mandated test. The open-ended items were scored by a mathematics educator and a high school teacher. The mean inter-rater correlations were .92 for the pre-test and .96 for the post-test. The content focus was addition of integers. "Levene's test of homogeneity of variance indicated that the variances of the two groups were not significantly different; $F(3, 96) = 0.16$, $p < 0.05$" (McClain & Berry, 2009, p. 461). Using a two-way factorial ANCOVA, results for the post-test indicated that the method of teaching had a statistically significant effect on achievement in favor of the teacher who engaged in culturally relevant pedagogy “$F(1, 96) = 13.06$, $p < 0.05$” (p. 466).

Although these results are not generalizable, the study suggests that when African American students experience teaching that is relevant to their lives, it can have a profound effect on achievement. On average, the students scored 6.5 points higher than the students taught by the traditional teacher.

Gutiérrez (2000) conducted a case study that looked beyond the classroom and individual teacher to examine how the culture of math departments may influence achievement. Utilizing the “Organized for Advancement Framework (OFA)” Gutiérrez investigated the following question: “How do the practices, beliefs, and general culture of a high school math department serve to encourage or discourage students from taking more math and higher levels of math?” (pp. 67-68). The OAF framework classifies an organization according to its potential to advance, on a continuum. The case study discussed here was part of a larger study that involved eight high school mathematics departments. Qualitative data included interviews, field notes, school documents and a questionnaire. These were examined through a process of creating “analytic case memos” through which prevailing themes emerged (p. 72). Using this information, the author developed scales to determine the level of success for each mathematics department. Quantitative data included six years worth of data on student course taking habits and mathematics achievement levels.

The high school that was the focus of the single case study was located in Maryland and had a student population categorized as low SES. Approximately 1,300 students attended the school with an African American student population slightly larger than 50%. The high school had educational resources that were comparable to similar schools.

> An examination of six years worth of quantitative data showed that: "…students at the school tend to take more math and higher levels of math than their counterparts in other high schools. They also maintain greater than predicted scores on mathematics achievement tests than their counterparts and are reported by their teachers to attend college at significant rates. Moreover, 22 of the 26 students in calculus are African American. (Gutiérrez, 2000, p. 100)

The author examined the qualitative data in order to discern the salient features of the mathematics department that contributed to the students' success. The school had adopted a rigorous mathematics curriculum with college prep courses ranging from Algebra 1 to Calculus. At the time of the study, it also offered business and consumer math, but these were being phased out during the year the study took place. Students were required to complete three mathematics courses in order to graduate. Classes contained students with mixed ability levels. Furthermore, there was a support system in place to keep students on track. In addition to regular coursework, a math lab was available to students who were less than proficient in basic skills. This allowed content teachers to review less. Teachers were also accessible to students before and after school for additional help. Moreover, they had a strong commitment to their students, focusing on their strengths and believing that it was possible for all students to succeed. Free summer school was offered to students who did not pass a course in order to keep them on track with their peers.

Reformed instructional practices were mandatory throughout the district and teachers were observed for their level of reformed teaching. Although some were slow to adopt this method of teaching, eventually the teachers understood the value of group work and collaboration, especially for African American students. In fact, this worked to a teacher's advantage because they often had 35 students in one class. The teachers created teams where students who mastered the content could assist their peers who were having difficulty.

The mathematics department chair played a significant role in student success. Rather than using her position to dictate how the math department was run, she acted as a support and resource for the teachers, helping teachers prepare rich activities that included the life experiences of the students. She made it a point to speak to each teacher every day. She also had a democratic style of assigning classes to teachers. She knew which courses teachers preferred to teach and gave them the opportunity to teach them, but she also rotated courses to ensure that all teachers had strong content knowledge in all courses. By rotating teachers across courses, they often taught the same students through different years, which helped to develop strong relationships.

The use of technology was widespread. Graphing calculators and computer software that was aligned to the curriculum were used often in problem-solving. The teachers also believed that while technology is a useful tool, it is not a substitute for number sense, and developing number sense was attended to in all courses. Finally, there was an atmosphere of "collective enterprise" (Gutiérrez, 2000, p. 84). The teachers relied on each other and shared resources. As teachers taught new courses, the teachers who previously taught the courses assisted them in preparation.

This case study tells the story of one high school and it is not generalizable, but it is important because it illustrates that collaboration among teachers and commitment to student success, along with challenging courses, reformed teaching practices, and high expectations can produce sustained improvements in achievement for African American students.

Balfanz and Byrnes (2006) studied three middle schools in the city of Philadelphia that had made great strides in improving achievement in mathematics. In general, students in the School District of Philadelphia begin middle school behind grade level and leave middle school even farther behind. The three schools studied were part of the Talent Development Middle Schools (TDMS) reform program. Students attending these schools were closer to grade level and some were even above grade level upon completion of middle school. The TDMS reform model includes the following components: mixed ability grouping; a challenging curriculum; standards-based and culturally relevant instruction; enrichment and extra support for students who need it; and, family and community involvement.

Students attending all three schools were primarily from low-income families. The majority of students were African American; one school had a large Latino/a population. Attendance rates for the schools ranged from 87% to 90%. Four cohorts of students from two schools and three cohorts from one school were examined. The students completed 8th-grade during the years 1999-2000 through 2002-2003. A total of 1,233 students comprised the sample. A student was included in the sample if the following information was available: “data for their pre-test, post-test (on the Stanford-9 Achievement Test; SAT-9), gender, race, school, cohort, effort (survey response), behavior, and attendance” (Balfanz & Byrnes, 2006, p. 148).

Using Logistic Regression Analysis, the authors investigated “the impact of certain enabling and constraining factors in whether or not students were able to close achievement gaps during their middle-school years" (Balfanz & Byrnes, 2006, p. 147). Several variables were controlled for in the analysis:

> students' average attendance rate over the time span, their average behavior marks in classes, their ethnicity, what cohort they were in, which of the three schools they attended, their effort in Math class…and finally, what percentage of their homerooms during the time span were 'High Gain Homerooms'….Students' gender was also available but was the only factor not to prove significant. (pp. 146-147)

All students attending the three TDMS schools were compared to 23 other Philadelphia middle schools with at least 80% minority students coming from low-income families. A t-test revealed a significant difference in the students' ability to catch up. Results of the SAT-9 indicated that 27% of the TDMS schools were catching up compared to 19% in the comparison schools (p = 0.000, t = 7.311, df = 13,696). When PSSA scores were examined, on average "TDMS students gained more than 10 state percentiles, while only 24% of students in the district's other schools gained as much…(p = 0.000, t = 8.167, df = 13,003)" (Balfanz & Byrnes, 2006, p. 150). When the regression for the sample was examined, it was determined that "an average of 42% of students [was] catching up on the SAT-9" (p. 150). The authors found that no one factor was predictive of catching up, but that in combination, high gain classrooms, attendance, effort, and behavior were predictive. This study illustrates that a reform model such as the TDMS model, when comprehensively implemented and sustained, can help to improve achievement in mathematics for students who often fall behind and fail to catch up.

Improving student achievement and opportunity to learn will continue to be at the heart of educational reform. The aforementioned studies are only a small sampling of successful attempts to narrow the achievement gap. The passage of NCLB has created an atmosphere of accountability. It is no coincidence that the launch of the Math Science Partnership Program occurred shortly after the passage of NLCB. Although the federal government holds states accountable, it has also increased available funding with the intent of increasing achievement and eliminating opportunity deficits that pervade much of K-12 mathematics and science education. However, the extent to which the MSP program has been able to accomplish this is yet to be determined.

Math and Science Partnerships: The Federal Government's Remedy for STEM Education

During September, 2002 the National Science Foundation (NSF) and the Department of Education announced 24 awards under a new Math and Science Partnership program. "MSP projects are expected to raise the achievement levels of all students and significantly reduce achievement gaps in the mathematics and science performance of diverse student populations" (NSF, 2009, p. 2). The MSP program is not unique in trying to accomplish this goal. In fact, these sentiments have been echoed nationally by educators, business and industry, and the government for many years. What is unique about the program is the extensive involvement of disciplinary mathematics and science faculty with K-12 education in pursuing this goal (NSF, 2006, p. 7). Silverstein, Frechtling and Miyaoka (2005) found that 73.5% of the lead partners for the first two cohorts of MSPs were either a university or a consortium of universities.

Dimitrov (2008) completed an initial analysis of student achievement data for three MSP cohorts: 2002-03, 2003-04 and 2004-05. The purpose of the investigation was to answer the following questions:

> What are the trends in mathematics and science proficiency changes across the entire three-year period for MSP-related schools that were previously investigated (Dimitrov, 2005) for such changes across the first two years (2002–03, 2003–04)? RQ2: What is the frequency distribution of MSP-related schools across categories of change (increase, decrease, or no change) in math and science proficiency and what is the mean effect size for the categories of significant change (increase or decrease) from the first (2002–03) to the third (2004–05) year? RQ3: What is the relationship between a school's targeted teacher participation in MSP related activities over the three-year period and the school's success in math and science proficiency at the end year of this period (2004–05)? (p. 638)

For the purpose of this review, only mathematics achievement will be discussed. Using the MSP management information system (MIS) data, Dimitrov found that on the elementary level the percent of students who achieved at least proficiency rose from 62.2% to 69.4% from 2003 to 2004 and up to 75.9% during the following year. These increases were statistically significant, $p < .05$. On the middle school level, students who achieved at least proficiency rose from 52.3% to 56.8%, to 58.5%. Again both increases were statistically significant, $p < .05$. High school students who were at least proficient rose from 42.1% to 58.7% ($p < .05$) during 2004 and then to 59.2% during 2005; these increases were not significant.

In order to answer the second question, proficiency was examined from year 1 to year 3. "The School's performance was labeled: (a) *increase*, if there was a statistically significant positive change; (b) *decrease,* if there was a statistically significant negative change; or (c) *no change*, if there was no statistically significant change for the school" (Dimitrov, 2008, p. 648). Applying Cohen's effect size index, Dimitrov found that schools that had a statistically significant increase also had the strongest effect sizes. The results for elementary schools show that "there are 71 schools with a positive change (medium mean effect size of 0.45) versus seven schools with a negative change (small mean effect size of 0.18)" (p. 649). For middle schools the comparison is nearly balanced with the effect size hovering around .19 for both increases and decreases. The most dramatic comparison occurred for high schools with a high effect size of .76 for increases and a low effect size of .09 for decreases. Question three investigated the relationship between teacher participation in professional development and increases in student achievement. Strong correlations were found in mathematics for elementary and secondary schools, $p < .01$, but no correlation was found on the middle school level.

During 2009, the RAND Corporation completed an interim evaluation on the Math Science Partnership of Southwest Pennsylvania. The partnership is funded through 2010. The monograph presents findings on how well the project is "meeting its three goals of increasing the mathematics and science achievement of K–12 students, increasing the quality of the K–16 educator workforce, and creating sustainable partnerships" (Pane, Williams, Olmsted, Yuan, Spindler, & Slaughter et al., 2009, p. 4). The evaluation team developed a logic model to assist in the analyses. The reader is cautioned that the goals stated above are long term goals. This monograph presents interim findings, on short-term and mid-term goals; therefore, final conclusions are not possible at this time. The project continues to collect quantitative and qualitative data and will evaluate whether the long term goals were met in the future. In spite of this, the monograph may inform other MSPs that are currently engaged in research.

The data sources consisted of mathematics standardized test data from 2000-01 through 2006-07 and science standardized test data for the years 2003-04 through 2006-07 for K-12. As well, a baseline survey was administered to all principals and a stratified sample of mathematics and science teachers during 2004. Follow-up surveys were administered to principals during 2006 and to teachers during 2007. Qualitative data were collected at the IHE level. Fifty-six "IHE-related faculty members and administrators" were interviewed using semi-structured interview protocols (Pane et al., 2009, p. 17). Some were interviewed more than once.

Results for increases in student achievement fall within mid-term goals. The team found an increasing trend through 2006-07. Given that it was not possible to attribute the increases explicitly to MSP involvement, the team examined whether there was a difference between MSP and non-MSP districts. The comparison non-MSP group included all school districts in Pennsylvania with the exception of the school districts of Philadelphia and Pittsburg, districts that were partners in the Math Science Partnership of Greater Philadelphia, as well as school districts that were not partners in an MSP, yet engaged in MSP activities. These districts were then matched on several variables and grouped. "After creating the groups, balance was assessed on all of the variables used in matching. Most balanced well, and those that did not were included as covariates in subsequent statistical models" (Pane et al., 2009, p. 25). The team found that non-MSP districts experienced similar rates of growth; however, the mean MSP proficiency scores were slightly lower than non-MSP proficiency scores. Next the team attempted to discern whether there was a relationship between achievement and participation in the MSP. Participation was dichotomized to include individual educator participation and district level participation. Due to difficulties in linking individual students with teachers, changes in student achievement levels were examined for only a small sample of students.

> A linear regression model was used to predict district level aggregated math achievement according to several factors: prior district-level aggregated math and reading achievement; the ratios of female, minority, and economically disadvantaged students; district-level leaders' total participation; and adjusted teachers' participation. (Pane et al., 2009, p. 24).

The evaluation team set the groundwork for investigating the two remaining goals: "increasing the quality of the K–16 educator workforce, and creating sustainable partnerships". However, these findings will not be discussed here as they relate to long-term goals (Pane et al., 2009, p. 4).

Kramer and Keller (2008) investigated whether the simultaneous implementation of block scheduling and a standards based reformed curriculum, the Interactive Mathematics Program (IMP), had any effect on student achievement and course taking in a suburban high school that was an MSPGP partner district. Both implementations were scaled up over the course of four years, beginning with 9th-grade students.

During the first year of implementation for 9th-graders, baseline scores were gathered for 11th-grade students who were not exposed to either innovation. Data included scores on "an Algebra Achievement Test (AAT) designed to measure both symbol manipulation emphasized by some more traditional texts and problem-solving emphasized by *Standards*-based curricula" (p. 4). During year two, the same students, who were now in twelfth grade, were evaluated using released items from the National Assessment of Educational Progress (NAEP). These evaluations were repeated on the following cohort, which also had no exposure to either block scheduling or a reformed curriculum. These groups served as the comparison groups for the first two cohorts of students who received both treatments throughout their entire high school experience.

Results showed the following statistically significant differences between the treatment and comparison groups ($p < .05$). When transcripts were compared, on average, the treatment groups spent more hours (577.5) enrolled in mathematics courses than did the comparison groups (498.0). However, the reader is cautioned that graduation requirements increased from three courses to four courses for treatment groups when compared to the comparison groups. Moreover, block scheduling by design allows students to complete more mathematics courses. When advanced course taking was examined, it was found that 55.1% of the treatment groups completed at least one advanced course compared to 47.8% for the comparison groups. Moreover, the percent of treatment students completing three or more advanced courses was 14.3% compared to 0.2% for the comparison groups.

When results for the AAT were compared, it was determined that the treatment groups scored higher on items which required students to: develop mathematical models ($d = .38$); translate "an exponential function from a table to an equation ($d = .67$)"; determine a rate of change from a graph ($d = 45$); identify functions to describe graphs ($d = .48$); and find "roots and the maximum of a quadratic ($d = .48$)". Comparison groups "were stronger at translating the graph of a linear function into an equation for that function ($d = –.48$), and at simplifying expressions with positive integer coefficients ($d = –.54$)" (Kramer & Keller, 2008, p. 5).

A comparison of the NAEP items yielded the following results. The treatment groups scored significantly higher than the comparison groups in all of the content areas measured by NAEP items: "Data Analysis, Statistics, and Probability, ($d = .40$); Geometry and Measurement, ($d = .40$); and Algebra and Number, ($d = .32$)" (p. 5).

Likewise, the treatment groups scored significantly higher on items requiring problem-solving (*d* = .36) and items requiring conceptual knowledge (*d* = .42).

These results are not generalizable to the larger population. This study was completed in one affluent high school. However, it is important because it suggests that the implementation of these two innovations, block scheduling and standards-based reformed curriculum, may in combination have a positive effect on both achievement and course completion.

Project PRIME (Promoting Reflective Inquiry in Mathematics Education), is an MSP located in South Dakota. The Rapid City Area School District, Black Hills State University, and "a regional education service agency" are the partners involved in the project (Sayler & Apaza, 2006, p. 2). The student population served by the Rapid City Area School District included approximately 36% who are eligible for free or reduced lunch and 17% who are Native Americans.

During its fifth year (2006) PRIME examined student data to determine if there were any changes in achievement since the project began. Data sources included the Dakota State Test of Educational Progress (DSTEP), a multiple choice test, and the "Balanced Assessments in Mathematics, developed by the Mathematics Assessment Resource Service and referred to hereafter as MARS tests…a free-response item test" (Sayler & Apaza, 2006, p. 4). The project also administered a baseline attitudinal survey during 2006, the *Attitudes Toward Mathematics Inventory* (ATMI) to establish baseline levels for grades six through twelve and to compare differences in attitudes between white and Native American students. Results showed that the mean score was higher during 2006 than during 2003 on the DSTEP for grades three through five by approximately 12%, and by nearly 20% for grades six through eight. For eleventh grade however, the mean score increased during the first year by approximately 2% and then declined through 2006, losing approximately 10% overall. When these results were compared to statewide results, the patterns were similar. When results were disaggregated, it was found that there was an achievement gap of approximately 30% between Native American and white students in PRIME schools. Again, this was consistent with state trends.

Results on the MARS for a random sample of students showed an increase of nearly 30% for grade four, an increase of approximately 12% for grade eight, and an increase for grade eleven of approximately 3%. It was also found that “the correlation between student-level test results on the two measures is highly significant with $p < 0.001$” (Sayler & Apaza, 2006, p. 18).

When results on the ATMI were compared for white and Native American students, it was found that the mean scores for white students were higher in confidence, value, enjoyment, and motivation. Conversely, Native American students’ mean scores were higher on items which suggest traditional teaching practices are more beneficial, such as, believing that mathematics “mainly involves memorizing and rules” and that students “learn more from listening than doing” (p. 22).

PRIME also collected data which measured the extent to which the project engaged in the following key elements: “1) district-wide graduate-level coursework for teachers; 2) building-based classroom coaching; 3) principal training; 4) parent nights; and 5) newly adopted inquiry-oriented instructional materials” (Sayler & Apaza, 2006,

p. 3). However, at the time of this report, it was not possible to determine whether there was any relationship between any key element and changes in student achievement. This study is important because in the future PRIME will investigate whether changes in student achievement and attitudes have any relationship with individual teachers, principals, or schools. Moreover, the project is poised to investigate cultural differences that may influence success in mathematics for Native American students.

Summary

It is evident from this review that reform efforts continue to be at the forefront of the Nation's educational mission. NCLB is but one remedy to improve student learning. "The global society is increasingly complex, requiring educated citizens who can learn continuously, and who can work with diversity, locally and internationally" (Fullan, 2007, p. 6). The literature also reveals that many reform efforts have failed to accomplish these ambitions goals. There are many pretexts, but three that seem to resonate throughout the literature are: organizational capacity, sustainability, and scaling up efforts. The body of literature examining the extent to which Math Science Partnerships have affected K-16 mathematics and science education is only beginning to emerge. While the preliminary findings for the most part appear to be positive in the area of student achievement, it is unclear whether these increases will persist. Therefore, the current study will contribute to the body of knowledge which documents and examines changes that occur in student outcomes in high school mathematics during the time of participation in a Math Science Partnership.

CHAPTER 3

METHODOLGY

Research Design

The purpose of this study is to determine whether involvement with an NSF funded Math Science Partnership (MSP) had any relationship to changes in mathematics performance for partner district high schools in Pennsylvania. The research questions investigated are:

1. What relationship, if any, exists between changes in student outcomes, on both the aggregate and disaggregate levels, in secondary mathematics and the level of participation by partner high schools in the Math Science Partnership of Greater Philadelphia?
2. Are there school district factors outside of participation in the Math Science Partnership that may have a relationship to changes in mathematics outcomes for students attending partner high schools?

This correlational study employs a quantitative design utilizing a modified version of cohort analysis. The data were not collected at the individual student level; therefore, cohort analysis is appropriate because it allows the researcher to…"examine the category as a whole for important features" (Neuman, 2002, p. 32). Because the data were collected over a 3-year period…"it is possible to identify when students change enrollment status….These datapoints allow…[the researcher]…to see the pattern of student progression for particular groups of students" (Voorhees & Lee, 2009, pp. 1 & 2). The purpose of the design is to look for the relationship between the level of involvement that the mathematics departments had in the MSPGP and the changes in student outcomes in high school mathematics. The study utilized a purposive sample and involved secondary data analysis. The data were not collected to answer the research questions; rather, the research questions were generated after an examination of the data that were available. Great care was taken in organizing the data (MSPGP staff members were consulted extensively and schools were contacted to verify information); however, the possibility exists that some data points are missing.

Analytical Framework

Activity Theory, the theoretical perspective that guides the study, will also serve as the analytical framework that will direct the analyses. The expanded model of Activity Theory allows one to view the MSPGP as a dynamic partnership which changed and developed in response to the needs and levels of commitment of its partners (see Figure 3-1). In this fluid system, Activity Theory "includes the actor (subject) or actors (subgroups)…[the] (object) as well as the dynamic relations among both" (Barab, Barnett, Yamagata-Lynch, Squire, & Keating, 2002, p. 78). Activity Theory is often situated under the umbrella of sociocultural theory because of the interrelatedness and interdependence of its features (Bandura, 1997; Engestrom, 1987; Seaman, 2008). According to the Center for Activity Theory and Development Work Research:

> In the model, the subject refers to the individual or sub-group whose agency is chosen as the point of view in the analysis. The object refers to the "raw material" or "problem space" at which the activity is directed and which is molded and transformed into outcomes with the help of physical and symbolic, external and internal mediating instruments, including both tools and signs [artifacts]. The community comprises multiple individuals and/or sub-groups who share the same general object and who construct themselves as distinct from other communities. The division of labor refers to both the horizontal division of tasks between the members of the

> community and to the vertical division of power and status. Finally, the rules refer to the explicit and implicit regulations, norms and conventions that constrain actions and interactions within the activity system. (University of Helsinki, n.d.)

Activity Theory focuses not on the individual per se but instead on the interactions between the subject and the components of the system (see Figure 3.1). According to Leontj'ev (1978), Activity Theory operates on three levels; activity, action, and operation. Each is oriented towards something and carried out by some agent. Activity is oriented towards an object or motive and carried out by the community. Action is oriented towards a goal and is carried out by an individual or a group. Operation is oriented towards conditions and is carried out routinely by humans or machines. "The activity system is a flexible unit of analysis (theoretical lens) which allows us to train our gaze in different directions and with different levels of 'magnification' to help us answer the questions that puzzle us" (Russel, 2002, p. 67).

The subject of the current study is the Math Science Partnership of Greater Philadelphia (MSPGP). A key goal of the project was to increase participation in challenging and higher level college preparatory mathematics courses which would lead to improvement in achievement (Outcomes). One way the partnership sought to accomplish this was to offer teachers and administrators various types of professional development that would promote reformed teaching practices (Tools). The breadth and depth of involvement to a large degree was mediated through the existing cultures within each school. Culture encompasses: the policies and procedures in place (Rules and Norms); the power structure that existed within schools (Division of Labor); and, the communities of practice that either existed prior to participation or resulted as a consequence of the participation. For example, the MSPGP acted as the agent to accomplish the goals, yet through its agency, became a member of many school district communities, developing relationships with administrators and working closely with teachers. The double arrows represent the fluid exchange of activity. However, in view of the fact that this study involves secondary data analysis, it does not capture every element that comprises the MSPGP activity system.

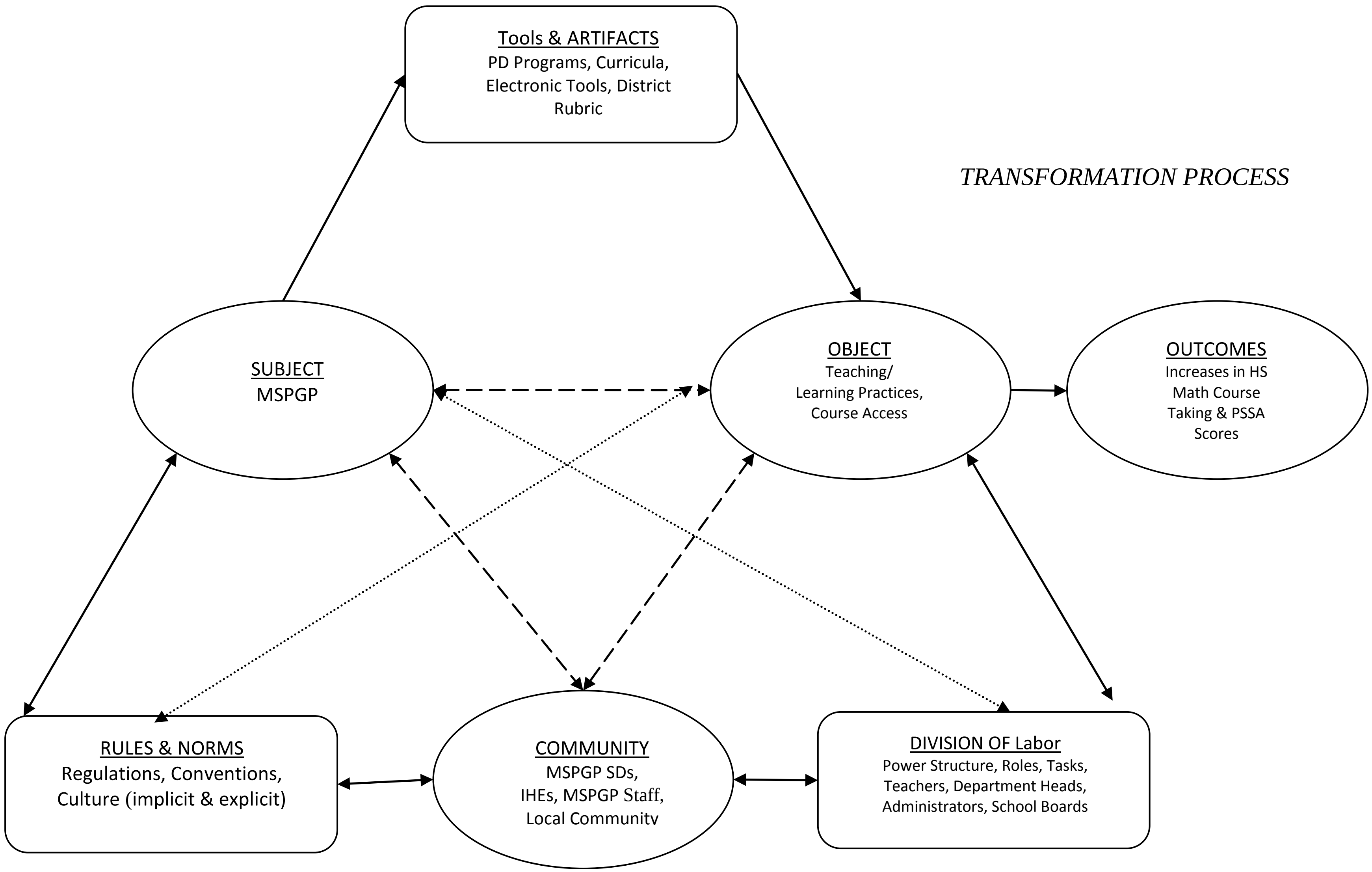


Figure 3.1. The MSPGP Activity System

Participants and Sampling

MSPs fall under the following categories: "Comprehensive, Targeted Institute Partnerships; Teacher Institutes for the 21st Century; and Research, Evaluation and Technical Assistance (RETA) projects". The Math Science Partnership of Greater Philadelphia (MSPGP) was a targeted partnership. Targeted partnerships focus "on improved K-12 student achievement in a narrower grade range or disciplinary focus within mathematics or science" (NSF, 2003, n. p.).

The MSPGP was a 5-year funded project that began in the fall of 2003. Its partners included 46 school districts and 13 institutes of higher education as well as an educational science center. Thirty-two school districts are located in the Greater Philadelphia region and 13 school districts are located in New Jersey. One charter school was also a partner. There are two community colleges, five 4-year colleges, and six universities that constituted the higher education partners. All IHE partners as well as the science center are located in Pennsylvania. The lead IHE partner was a university located in Philadelphia. The MSPGP focuses on both mathematics and science education in grades 6 through 12.

This study utilized a purposive sample. The sample was limited to 23 Pennsylvania school districts that were MSPGP partners for which the project had complete data sets. The project did not have complete data sets for nine of the remaining 10 Pennsylvania school districts and one district was eliminated because its course completion data skewed the distribution. This high school required six math courses for graduation, while most of the other Pennsylvania districts required only three or four. The

New Jersey districts were not considered in the study because the math standards and standardized testing practices in New Jersey differ from that of Pennsylvania. The New Jersey High School Performance Assessment (HSPA) is an exit exam, while the PSSA is not. Moreover, students in New Jersey "are first tested…during the spring semester and, if they fail…again during the fall and/or spring semesters in 12th grade" (Ou, 2009, p. 4). Pennsylvania allows retesting but does not require a passing score to graduate (PA department of Education). Therefore, a cross state analysis may confound the results because State Education Agencies (SEAs) have different organizational structures and directives for school districts. At the high school level in the sample districts, the percentage of African American and Latino/a populations combined ranged from 1.1% to 76.1%, with a median of 8.8%. The percentage of students who were eligible for free or reduced lunch ranged from 1% to 66% with a median of 11.5%. The percentage of students identified as low socio-economic status (SES) ranged from 1.8% to 60.5% with a median of 9.65%. The grade levels under investigation are ninth through twelfth. Pseudonyms are used for all of the high schools in this investigation (see Table 3.1).

Table 3.1. MSPGP Pennsylvania High School: African American and Latino/a Populations

District	African American	Latino/a	Total
Jericho	0.3%	0.8%	1.1%
Carter	1.1%	0.7%	1.8%
Fairlea	1.6%	1.5%	3.1%
Buchanan	1.9%	2.8%	4.7%
Twain	1.8%	3.0%	4.8%
Roosevelt	1.9%	3.4%	5.3%
Midvale	2.6%	3.1%	5.7%
Taylor	5.5%	0.9%	6.4%
Grant	2.1%	5.0%	7.1%
Taft	5.9%	2.3%	8.2%
Rye	6.9%	1.5%	8.4%
South Harding	7.8%	1.0%	8.8%
Hayes	7.4%	2.7%	10.1%
Jackson	4.4%	6.5%	10.9%
Erikson	9.2%	1.9%	11.1%
Cambridge	9.9%	1.2%	11.1%
Rittenhouse	17.9%	2.2%	20.1%
Carver	13.8%	7.9%	21.7%
Tyler	19.3%	14.9%	34.2%
Madison	9.5%	30.2%	39.7%
North Ford	43.9%	3.3%	47.2%
Anthony City	17.7%	55.4%	73.1%
Pierce	24.2%	51.9%	76.1%

Data Sources

The dependent or outcome variables examined were changes in mathematics performance. The independent variables examined were: publically available data and MSPGP data that were collected during the time that the project was in existence. Once the research questions were generated, the data were further examined to check for validity and reliability. This was done through a careful comparison of the electronic data sources and the hard copies that originally documented project activities, as well as through lengthy discussions with the project staff who were involved in the data collection. Because the data were not collected to answer a research question or test a theory, some limitations exist which will undoubtedly affect the depth and breadth of the inquiry, as well as the conclusions drawn. Where the data are deemed unreliable, the limitations that arise will be discussed extensively. This discussion is important for framing the context of an analysis on the partnership level and for informing the next generation of MSPs that are currently planning or engaging in research and evaluation about the impact of their projects.

Public data sources included data accessed through the Pennsylvania Department of Education website. MSPGP data sources included information collected on: high school mathematics course taking patterns; teacher professional development; and the level of engagement between school districts and the MSPGP. The grant years under investigation were 2005-06 (year 3) through 2007-08 (year 5). Mathematics course completion data were collected during 2003-04 (year 1) via a paper survey; however, much of the year 1 data were unreliable. This was due, in part, to the challenges imposed on districts to assemble data in a format that was unfamiliar to many of them. Therefore, any analysis based on year 1 data would likely be invalid. Hence, it is not considered in this investigation. No data were collected during 2004-05 (year 2) due to the difficulties encountered during the first data collection. By the third year these data collection challenges were overcome through intense communication with school districts and the development of an online data collection mechanism that had validation capabilities.

Data points were examined and grouped into categories. The outcome or dependent variables are: changes in high school level college prep mathematics course completion; changes in performance on the mathematics portion of the 11th-grade PSSA; and, changes in achievement gaps for underrepresented student populations. Course completions were captured through the online data collection survey described above. The survey compiles course completions disaggregated by: course level (e. g., level 1 may be Algebra 1 or integrated mathematics 1), grade level, gender, and race/ethnicity. These raw data were converted to percentages through an electronic process which generated excel spreadsheets (see Appendix A for a sample table). The system was not programmed to disaggregate by gender until year 4; therefore, for the purpose of this study, gender will not be considered in the analyses.

PSSA data were disaggregated by: gender (Male/Female), race/ethnicity (White, Black[1], Hispanic[2], Asian, Native American, and Multi-ethnic), Individualized Education Plan (IEP), English Language Learners (ELL), and Economically Disadvantaged. Students are scored on a proficiency basis. The levels of proficiency are: Below Basic, Basic, Proficient, and Advanced. Reliabilities for the 11th-grade mathematics PSSA ranged from 0.92 to 0.94 for the years 2002 and 2003 (Mead, Smith, & Swanlund, 2003). During 2004, the Human Resources Research Organization (HumRRO) investigated the correlation between the math PSSA and math SAT scores for 2002 and 2003. Correlations were high ($r = 0.846$ and $r = 0.865$, respectively) for both years (Koger,

[1] The identifier Black is used instead of African American to match the terminology used in the PSSA.
[2] The identifier Hispanic is used instead of Latina/o to match the terminology used in the PSSA.

Thacker, & Dickson, April, 2004). For the purpose of this study, aggregated PSSA data and PSSA data disaggregated by race/ethnicity are considered.

Through careful examination of the independent variables, two categories of school level data emerged: teacher professional development, and district stability. Teacher professional development variables included the type of the professional development attended as well as the hours accrued by each participating teacher. Professional development topics included: content, curriculum (curriculum alignment, curriculum development, special training in curriculum), data analysis, leadership (goal setting, visioning), pedagogy, technology, and other non-specified professional development. These data were housed in the professional development database that was created by the project to track this information. This database was recently reviewed for accuracy by several project staff members who organized, scheduled, and tracked professional development for the project. Because much of the teacher training took place during the early years of the grant, and because the effects of professional development may not be realized until well after it is received (Guskey, 1997; 2002; Snow-Renner & Lauer, 2005), the total amount of professional development accrued over the five years of the project will be included in the analyses. For the purpose of this investigation, professional development in aggregate form will be considered in the analysis. That is, the investigation will not evaluate the effect that each type of professional development may have had. Rather, it will consider professional development in its entirety.

School engagement occurred on three levels: engagement with IHE faculty; engagement with MSPGP staff; and engagement in professional development. The level of engagement between IHE faculty and school districts was tracked through the MSP Management Information System (MIS). Each year, partner IHE faculty completed a survey that described the types of activities they engaged in, such as research or creating college level or high school level courses, and the number of hours dedicated to this work. However, the system does not document the number of contact hours IHE faculty dedicated specifically to individual schools; therefore, this form of IHE engagement is not considered in the analyses. Faculty members also organized and ran professional development sessions specifically for individual schools. This was tracked through the professional development database and will be considered in the analyses; however, conclusions will be drawn with caution because the level of engagement documented through the professional development database may not adequately represent the actual level of engagement between higher education and school district partners.

MSPGP engagement was documented through the creation of a School District Rubric, which evaluated where the district was located on a reform indicator continuum, as well as the level of engagement the district had with the partnership. The levels of reform included: Struggling, Status Quo, Getting Ready, On the Road, and High Achieving (see Appendices B & C). The levels of engagement were: None, Low, Moderate, High and Very High. For the purpose of this study, only the levels of engagement will be considered. Determining a method to quantify the level of reform was problematic. For example, according to the reform continuum definitions, schools that are status quo "either see no need to change any of their practices because their students are already fully challenged and are scoring well, or if their students are not performing optimally, they do not want to spend the time and resources necessary to change" (see Appendix C). This type of ambiguity does not easily allow for a reliable or valid indication of growth or commitment to reform; therefore, only the level of engagement with the MSPGP will be considered in the analyses.

District ratings of levels of engagement were determined by consensus by several project staff members who had ongoing contact with districts. However, this tool was not field tested, nor was inter-rater reliability established. This was due to the fact that the rubric was used as a tool to give a snapshot of the status of districts on the road to reform. It was not conceptualized as a data set to be used for analysis of the project. Therefore, the ratings assigned to schools were not validated. However, the rubrics include rich narrative descriptions as well as the ratings. These descriptions will be mapped to ratings to examine emergent themes.

School district stability variables include: the type of curriculum used, whether reformed or traditional; high school scheduling, whether traditional or block; the total minutes dedicated to teaching a full math course, assuming a 180-day academic year; and, administrator turnover, including a change in the superintendent or principal. The findings on systemic initiatives alluded to earlier in the literature review suggest that change is a slow process and may take longer than five years; therefore, it is important to examine whether or not factors external to participation in the MSPGP had any influence on changes in outcomes.

Procedures

In order to answer the research questions, the expanded model of Activity Theory will be employed as the methodological tool to identify the unique features of the MSPGP that may have had an impact on mathematics high school performance levels (see Figure 3.1). Recall, that performance is defined as: proficiency levels on the PSSA; mathematics course completions; and, changes in achievement gaps. Comparisons will be made between 2005-06 (year 3) and 2007-08 (year 5).

The level of participation in professional development for teachers was generated through the professional development database (Tools & Artifacts). Participation is defined as the average number of professional development hours that mathematics teachers attended (see Table 3.2 for example; see Appendix D for complete table).

Table 3.2 Teacher Professional Development Hours (example)

School District	Academic Year	Total Number of Participating Teachers	Total Professional Development Hours	Average Professional Development Hours
Anthony City	2003-04	0	0.00	0.00
	2004-05	28	452.50	16.20
	2005-06	18	200.50	11.14
	2006-07	34	328.50	9.66
	2007-08	2	35.50	17.75
	TOTAL		1,017.00	54.71

School administrators also participated in professional development. Therefore, the mean number of hours in which administrators engaged in professional development for the duration of the project will represent administrator participation. IHE involvement in professional development is defined as the average number of hours partner faculty organized and presented professional development. Although many faculty members also attended professional development, it is not clear that attendance translates into engagement with schools. For example, a faculty member could have attended professional development to further his/her own knowledge or to further his/her work with teachers in partner high schools. This is not apparent from the data sources available; therefore, only the number of hours that IHE faculty led professional development will be considered in the analyses.

Professional development engagement is calculated as: $e = pd_1 + pd_2 + pd_3$, where pd_1 is equal to the average number of professional development hours accrued by high school mathematics teachers, pd_2 is equal to the average number of professional development hours accrued by administrators, and pd_3 is equal to the average number of professional development hours provided by IHE faculty (see Table 3.3).

Scatter plots were generated to explore whether there was a relationship between the total number of professional development hours and changes in: the percentage of students scoring at least proficient on the PSSA; the percentage of students completing college prep mathematics courses; and, the percentage of students completing higher level college prep mathematics courses. The graphs suggest weak but positive correlations among all three associations (see Appendix E). Therefore, continuing this line of inquiry is warranted to determine whether the level of professional development did in fact have any relationship to changes in outcomes.

Table 3.3. Overall Professional Development Engagement

School District	Teacher Attended (average)	Administrator Attended (average)	IHE Provided (average)	Total
Buchanan	8.25	0	5	13.25
Taft	9.88	10	1.5	21.38
Rye	20.81	7.75	0	28.56
North Ford	40.08	0	0	40.08
South Harding	31.5	0	18	49.50
Twain	26.93	9.75	19.5	56.18
Jackson	18.25	0	42	60.25
Midvale	43	3	27	73.00
Madison	68.94	0	6.5	75.44
Cambridge	77.41	0	15	92.41
Eriksson	55.42	5.13	43.25	103.80
Roosevelt	66.61	31.5	8	106.11
Tyler	83.8	0	23.25	107.05
Rittenhouse	62.88	2	50.25	115.13
Fairlea	61.38	24.67	30	116.05
Taylor	81.07	30.5	16.75	128.32
Pierce	66.92	62	0	128.92
Anthony City	54.75	24.67	52.46	131.88
Hayes	65.27	19	60	144.27
Jericho Township	108.5	0	36	144.50
Carter	135	0	27	162.00
Grant	109.75	0	60	169.75
Carver	119.34	18.38	70.83	208.55

Professional development was dichotomized into high and low (see Table 3.4). High and low indicators were determined by first generating a histogram of the data. The histogram revealed a clear delineation at 100.00 hours (see Appendix F). The mean was calculated, M = 98.97 and the median was determined to be 106.11. The mean of these three values was calculated, M = 101.69. Thus, the category of low was defined as less than 102 hours and the category of high was defined as greater than or equal to 102 hours.

Table 3.4. Level of Engagement: Professional Development Hours

Mean	Median	Histogram Delineation	Average	High	Low
98.97	106.11	100.00	101.69	≥ 102	< 102

The level of engagement that a school had with the MSPGP was based on two types of activities: the percent of eligible teachers who participated in professional development and the rating of engagement that was documented through the District Rubric.

The percent of eligible teachers who participated in professional development was calculated in the following way. The total number of eligible teachers was broken down for each year of the grant. This number was supplied by the school district on the district survey that collected course enrollment information. Data were not collected during the 2004-05 academic year. In order to determine this number, high schools were contacted via phone and asked how many teachers were in the math department during 2004-05. They were also asked to verify the number of math teachers for the remaining years for which the project had collected data. Once a valid eligibility value for each year of the project was determined, it was divided into the yearly total number of math teachers who participated in professional development. This quotient represents the percentage of eligible teachers who participated for any given year. Next, these quotients were averaged to determine an engagement value for the life of the project (see Table 3.5 for example; see Appendix D for complete table).

Table 3.5 Percent of Eligible Teachers Engaged in Professional Development

School District	Academic Year	Total Number of Eligible Teachers	Total Number of Participating Teachers	Percent of Eligible Teachers
Anthony City	2003-04	49	0	0%
	2004-05	47	28	60%
	2005-06	48	18	38%
	2006-07	47	34	72%
	2007-08	45	2	4%
	Percent Engaged			35%

The School District Rubric rating scale consisted of: None, Low, Moderate, High and Very High. These ratings were converted to a scale of zero through four. Ratings were summed across years to determine a score that represented the total level of engagement. The highest possible score was 20. A histogram of the data revealed a clear delineation at 7.5 (see Appendix F). The mean was calculated, *M* = 7.1, and the median was determined to be 8.0. The mean of these three values was calculated, *M* = 7.53 (see Table 3.6). Thus, the category of low was defined as a score of less than 7.5 and the category of high was defined as greater than or equal to a score of 7.5.

Table 3.6. Level of District Engagement with the MSPGP: School District Rubric

Mean	Median	Histogram Delineation	Average	High	Low
7.10	8.00	7.50	7.53	≥ 7.5	< 7.5

In order to try to minimize the threat to internal validity that existed due to the lack of inter-rater reliability, a keyword search of the words: leader, coach, reform, mentor, and formative assessment was completed on the full rubric, including the districts that were not part of this study. This was done in order to tease out the themes of partnership activities. The instances of appearance were: reform (97), leader (44), coach (38), formative assessment (32), and mentor (8). Although mentoring was encouraged in partnership activities, the word mentor was eliminated due to the few instances of its appearance in the narratives. Only instances of appearance that were indicative of positive growth were summed. For example ‘administration is committed to reform in the district’ was counted, while ‘the district is resistant to reform’ was not.

A second keyword search was completed on the sample districts and the instances of appearances were summed. This total was then compared with the engagement scores that were generated through the rubric. It was found that districts whose scores were eight or above had between seven and 16 appearances of the keywords, except for one district that had only six. Those with scores less than eight had between two and six appearances of the keywords. This partially supports the dichotomization of high and low engagement at a score of 7.5 (see Table 3.7).

Table 3.7. School District Rubric: Engagement Summary

Low Engagement

School District	*Rubric Sum*	Key Words: Instances of Appearances Coach	Leader	Reform	Formative Assessment	Total Keywords
Buchanan	5	2	1	1	1	5
Carter	4	0	2	1	0	3
Grant	5	1	1	2	1	5
Jackson	3	0	1	3	0	4
Midvale	3	0	1	1	0	2
Pierce	6	1	1	4	0	6
Rye	7	0	2	3	0	5
South Harding	3	0	0	2	0	2
Taft	5	0	1	2	0	3
Twain	5	1	1	2	0	4
Tyler	6	4	0	2	0	6

High Engagement

School District	*Rubric Sum*	Key Words: Instances of Appearances Coach	Leader	Reform	Formative Assessment	Total Keywords
Anthony City	8	4	2	7	1	14
Cambridge	8	0	2	3	8	13
Carver	10	0	2	4	0	6
Eriksson	13	8	1	4	1	14
Fairlea	8	2	3	2	0	7
Hayes	8	1	1	5	1	8
Jericho	8	0	3	4	0	7
Madison	8	2	2	6	0	10
North Ford	8	0	1	4	2	7
Raylor	13	3	0	3	10	16
Rittenhouse	10	0	0	9	0	9
Roosevelt	9	0	4	5	1	10

The district which had a total rubric score of 10 with only six appearances of the keywords had accrued the second highest number of professional development hours; therefore, although the number of keyword appearances was low, the level of professional development in which math teachers participated in, warrants placing it in the category of high.

The dichotomy is moderately supported through a significant correlation between the rubric sums and the percentages of eligible teachers, who participated in professional development, $r = .572$, $p = .004$. When separate correlations for the rubric sums representing high and low breakdowns were calculated, results show the correlation for schools in the high category was significant, $r = .611$, $p = .046$ as was the low category, $r = .677$, $p = .016$.

The unit of analysis is the high school. The dependent variables that will be examined are: changes in the percentage of students scoring at least proficient on the mathematics portion of the PSSA; changes in the percentage of students completing high school level college prep mathematics courses; changes in the percentage of students completing higher level courses; and, changes in achievement gaps between white and African American, and white and Latino/a students. The two independent constructs considered in the analyses are the level of participation in professional development, dichotomized as high and low, and engagement as it relates to the level of involvement between the MSPGP and the high school, including instructional leaders and administrators, also dichotomized as high and low. In order to answer the first research question, the biserial correlation will measure 30 possible relationships (see Table 3.8). The biserial correlation is warranted because it measures the relationship between a continuous variable and a dichotomous variable where the dichotomy does not occur naturally and where an interval or continuous variable underlies the dichotomy (Humphreys & Swets, 1991; Kline, 2005).

Table 3.8. MSPGP Inputs and Outputs

	OUTPUT		
INPUT	Changes in the Percentage of students scoring at Least Proficient on the PSSA	Changes in the Percentage of College Prep Math Courses Completed	Changes in the Percentage of Higher Level College Prep Math Courses Completed
Level of Professional Development			
	All Students African American Latino/a	All Students African American Latino/a	All Students African American Latino/a
	African American/White Achievement Gap	African American/White Achievement Gap	African American/White Achievement Gap
	Latino/a/White Achievement Gap	Latino/a/White Achievement Gap	Latino/a/White Achievement Gap
Level of Engagement with MSPGP			
	All Students African American Latino/a	All Students African American Latino/a	All Students African American Latino/a
	African American/White Achievement Gap	African American/White Achievement Gap	African American/White Achievement Gap
	Latino/a /White Achievement Gap	Latino/a/White Achievement Gap	Latino/a/White Achievement Gap

It may be the case that changes in outcomes were related to other independent variables; therefore, in order to answer question 2, Pearson correlations were run to determine whether changes in outcomes had a relationship with the components of the Activity System that are located at the bottom of the triangle: Rules and Norms, Community, and the Division of Labor. School district stability variables fall within these components. They may also have overlapping ties to multiple components. For example, a school district's decision to implement block scheduling may be encompassed in Rules and Norms but may also be a consequence of pressure from the community to offer more mathematics courses to students. The reader is cautioned that this list of school district variables is not exhaustive; there are many other school district variables that were not gathered by the project that may have influenced student outcomes.

Summary

Compiling the data in order to analyze how changes in student outcomes are linked to MSPGP activities was challenging. Before the analyses are completed and discussed, it is important to reiterate that these data do not provide a complete picture of partnership activities. Although great care was taken to organize the data in a reliable manner, limitations exist and will be discussed thoroughly in Chapter Four and Chapter Five.

CHAPTER 4

RESULTS

Introduction: Changes in Student Outcomes

The purpose of this study was to investigate whether participation in the Math Science Partnership of Greater Philadelphia (MSPGP) by high schools had any effect on secondary students' outcomes in mathematics. The intent of the study is to contribute to the growing body of literature examining student achievement and inequity in K-12 mathematics education. The results of the analyses will answer the following research questions:

1. What relationship, if any, exists between changes in student outcomes, on both the aggregate and disaggregate levels, in secondary mathematics and the level of participation by partner high schools in the Math Science Partnership of Greater Philadelphia?
2. Are there school district factors outside of participation in the Math Science Partnership that may have a relationship to changes in mathematics outcomes for students attending partner high schools?

The dependent variables considered were: changes in the percent of students scoring at least proficient on the PSSA; changes in overall college prep high school level mathematics course completions; changes in higher level course completions; and, changes in achievement gaps. The overarching goal of the MSP program is to "raise the achievement levels of all students and significantly reduce achievement gaps in the mathematics and science performance of diverse student populations" (NSF, 2002, n. p.).

In order to determine whether any changes occurred in mathematics outcomes, comparisons were made between grant year 3 (2005-06) and grant year 5 (2007-08). First, the data were examined to determine whether changes in each school were greater than would be expected by chance. Of the 23 high schools examined, 18 or 78% had an increase in the percent of students scoring at least proficient on the PSSA. A binomial test revealed that the probability that the number of schools that had increases was not simply due to chance and was significant, $z = 2.5$, $p = .011$. The mean percentage for the number of students scoring at least proficient on the PSSA for 2006 was 58.5%. This increased to 64.1% during 2008. A paired samples t-test revealed that this increase was significant, $t = 4.03$, $df = 22$, $p = .001$. This result is similar to the statewide trend in Pennsylvania. Sixteen or 70% of the schools had an increase in total college prep mathematics course completions and 15 or 65% of the schools had an increase in higher level course completions, courses above Algebra 2 or Integrated Math 3. Binomial tests revealed that these probabilities were not significant. The mean percentage for overall college prep course completions for 2006 was 81.6%. This increased to 85.8% during 2008. The mean percentage for 2006 for higher level college prep course completions was 22.1%. This increased to 26.5% during 2008. Paired samples t-tests revealed that neither increase was significant (see Table 4.1).

Table 4.1. Changes in Mathematics Outcomes: All Students Grades 9-12

School Districts	PSSA Advanced & Proficient Net Change (%)	CP Course Completion Net Change (%)	Higher Level CP Course Completion Net Change (%)
Anthony City	11	18	4
Buchanan	-3.4	10	4
Cambridge	3.3	2	5
Carter	7.7	16	13
Carver	8.6	-19	-20
Eriksson	11	25	21
Fairlea	-7.4	4	5
Grant	6.2	-4	-1
Hayes	9	17	-3
Jackson	11.5	-11	-4
Jericho Township	15.7	6	5
Madison	5.6	6	7
Midvale	1.9	-5	-7
North Ford	-3	-8	5
Pierce	1.7	4	4
Rittenhouse	11.8	9	29
Roosevelt	13.1	-8	1
Rye	11.5	-10	-7
South Harding	-0.6	9	5
Taft	1.4	4	-1
Taylor	11.4	2	23
Twain	-6.9	26	13
Tyler	6.3	2	0

When changes for African American students were examined, 13 or 93% of the 14 schools for which PSSA data were available showed an increase in the percent of students scoring at least proficient on the PSSA. A binomial test revealed that the number of schools that had increases was not simply due to chance and was significant, $z = 2.94$, $p = .003$. The mean percentage for the number of African American students scoring at least proficient on the PSSA for 2006 was 28.2%. This increased to 35.9% during 2008. A paired samples t-test revealed that this increase was significant, $t = 2.53$, $df = 13$, $p = .025$ (see Table 4.2). This result is similar to the statewide trend in Pennsylvania. Empty cells on the table indicate that either there were no African American students tested or the African American population of a school was too small to report.

Table 4.2. Changes in Mathematics Outcomes: African American Students Grades 9-12

School Districts	PSSA Advanced & Proficient Net Change (%)	CP Course Completion Net Change (%)	Higher Level CP Course Completion Net Change (%)
Anthony City	12.1	28	14
Buchanan		23	4
Cambridge	8	2	1
Carter			
Carver	10.7	-27	-18
Eriksson	26.3	19	13
Fairlea		4	-5
Grant		-13	-4
Hayes	-2.3	7	3
Jackson	22.7	-12	1
Jericho Township			
Madison	7	-1	2
Midvale		-14	5
North Ford	0.4	-13	6
Pierce	5.5	6	3
Rittenhouse	17.7	-10	5
Roosevelt			
Rye	6.6	-33	-8
South Harding	1.8	34	4
Taft		8	0
Taylor	14	1	1
Twain		38	3
Tyler	2.9	7	0

Twelve or 60% of the 20 schools that reported course completions for African American students had an increase in total college prep mathematics course completions and 14 or 70% of the schools had an increase in higher level course completions. Binomial tests revealed that these probabilities were not significant. The mean percentage for overall college prep course completions for African American students for 2006 was

66.8%. This increased to 69.5.8% for 2008. The mean percentage for higher level college prep course completions for African American students for 2006 was 10.6%. This increased to 12.1% during 2008. Paired samples t-tests revealed that neither increase was significant (see Table 4.2).

When changes for Latino/a students were examined, 7 or 87% of the 8 schools for which PSSA data were available showed an increase in the percent of students scoring at least proficient on the PSSA. Twelve or 60% of the 20 schools that reported course completions for Latino/a students had an increase in total college prep mathematics course completion. Binomial tests revealed that neither probability was significant. Seventeen or 85% of the schools had an increase in higher level course completion for Latino/a students. A binomial test revealed that the probability that the number of schools that had increases was not simply due to chance and was significant, $z = 2.91$, $p = .003$. The mean percentage for the number of Latino/a students scoring at least proficient on the PSSA for 2006 was 26.31. This increased to 37.2 during 2008. The mean percentage for overall course completions for Latino/a students for 2006 was 66.8%. This increased to 69.5% during 2008. Paired samples t-tests revealed that these increases were not significant. The mean percentage for higher level course completion for Latino/a students for 2006 was 66.5%. This increased to 73.3% during 2008. A paired samples t-test revealed that this increase was significant, $t = 3.08$, $df = 19$, $p = .006$ (see Table 4.3). Although the majority of improvements were not at the level of significance, most high schools experienced growth in outcomes, both overall as well as for African American and Latino/a students.

Table 4.3. Changes in Mathematics Outcomes: Latino/a Students Grades 9-12

School Districts	PSSA Advanced & Proficient Net Change (%)	CP Course Completion Net Change (%)	Higher Level CP Course Completion Net Change (%)
Anthony City	14.7	20	4
Buchanan		16	5
Cambridge		68	30
Carter			
Carver	34.9	-29	-16
Eriksson		0	29
Fairlea		12	3
Grant		-35	-4
Hayes	11.8	-9	4
Jackson	-9.9	-7	1
Jericho Township			
Madison	5.5	4	3
Midvale		3	10
North Ford	24.3	0	6
Pierce	3.6	6	5
Rittenhouse		27	22
Roosevelt			
Rye		12	6
South Harding		52	7
Taft		3	11
Taylor		-20	6
Twain		13	23
Tyler	4.4	-1	-2

A comparison of changes in achievement gaps follows. Comparisons were made between white students and African American students and between white students and Latino/a students. A negative value indicates that the achievement gap decreased while a positive value indicates that the achievement gap increased. The reader is cautioned in interpreting changes in achievement gaps. Although any decrease in achievement gaps is notable, population percentages vary among the 23 schools. A small decrease in achievement gaps for a school with large minority populations may be more impactful than a larger decrease for a school with fewer minority students.

Table 4.4. Changes in Achievement Gaps 2006 – 2008: African American Students

School Districts	African American Percent of High School Population	PSSA Net Change (%)	CP Course Completion Net Change (%)	Higher Level CP Course Completion Net Change (%)
Anthony City	17.7	0.6	-17	-20
Buchanan	1.9		-13	0
Cambridge	9.9	-5.5	-2	-15
Carter	1.1			
Carver	13.8	-5	8	-8
Eriksson	9.2	-17.1	7	-16
Fairlea	1.6		-1	-25
Grant	2.1		14	4
Hayes	7.4	13.1	7	8
Jackson	4.4	-9.4	0	-14
Jericho Township	0.3			
Madison	9.5	1.4	11	3
Midvale	2.6		9	-27
North Ford	43.9	-2.6	11	2
Pierce	24.2	-0.5	-5	4
Rittenhouse	17.9	-5.9	7	-12
Roosevelt	1.9			
Rye	6.9	6.6	23	-23
South Harding	7.8	-0.9	-27	4
Taft	5.9		-5	-3
Taylor	5.5	-2.3	3	1
Twain	1.8		-11	2
Tyler	19.3	5.6	-4	4

When comparisons were made between white and African American students, the achievement gap in PSSA proficiency decreased for nine or 64% of the schools for which PSSA data were available (see Table 4.4). The gap in college prep course completion decreased for nine or 45% of schools while the gap in higher level course completion decreased for 10 or 50% of the schools.

Table 4.5. Changes in Achievement Gaps 2006 – 2008: Latino/a Students

School Districts	Latino/a Percent of High School Population	PSSA Net Change (%)	CP Course Completion Net Change (%)	Higher Level CP Course Completion Net Change (%)
Anthony City	55.4	-2	-9	0
Buchanan	2.8		-6	-1
Cambridge	1.2		-68	-14
Carter	0.7			
Carver	7.9	-29.2	10	-4
Eriksson	1.9		26	-9
Fairlea	1.5		-9	1
Grant	5		36	4
Hayes	2.7	-1	23	-1
Jackson	6.5	23.2	-5	-3
Jericho Township	0.8			
Madison	30.2	2.9	6	6
Midvale	3.1		-8	-17
North Ford	3.3	-26.5	-2	3
Pierce	51.9	1.4	-5	1
Rittenhouse	2.2		-30	-3
Roosevelt	3.4			
Rye	1.5		-22	-10
South Harding	1		-45	-2
Taft	2.3		0	-10
Taylor	0.9		24	17
Twain	3		14	-10
Tyler	14.9	4.1	4	4

When comparisons were made between white and Latino/a students, the achievement gap in PSSA proficiency decreased for four or 50% of the schools for which PSSA data were available (see Table 4.5). The gap in college prep course completion decreased for 11 or 55% of schools, while the gap in higher level course completion decreased for 12 or 60% of the schools.

Summary

The purpose of the preceding section was to describe the changes that occurred in student outcomes for MSPGP partner high schools in mathematics. This is illustrated in the central region of the expanded model of the activity system and as such is an incomplete picture of the MSPGP activity system (see Figure 4.1). Because results were mixed, that is, some schools experienced increases in some outcomes, while other schools experienced decreases in some outcomes, one cannot conclude that participation alone in the MSPGP brought about changes. There are other factors which may have contributed to changes in outcomes that are worth exploring. What follows is an examination of the relationship between the characteristics of MSMGP engagement and changes in outcomes to determine if any relationship did exist.

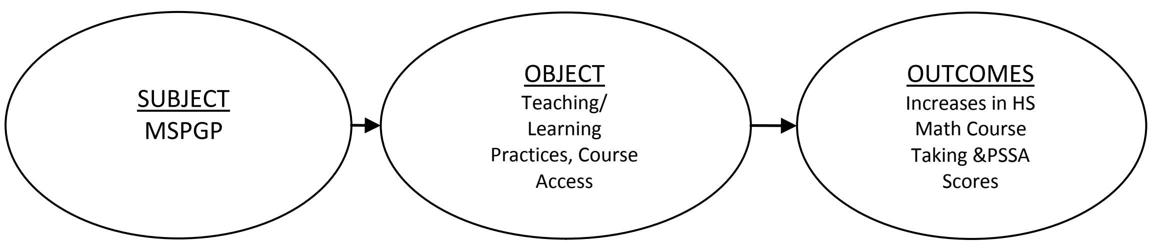


Figure 4.1. Activity Theory: Expanded Model: Central Region

Correlations: Changes in Outcomes and Participation in the MSPGP

The biserial correlation was calculated to investigate whether changes in outcomes had any relationship with the level of participation that high school math teachers had with the MSPGP (see Table 4.6). SPSS does not have a biserial correlation option; however, the statistic can be generated using the point-biserial correlation coefficient, a capability that SPSS does possess (Field, 2009). The formula is:

$$r_p = (r_{pb} \sqrt{P_1 P_2}) / y$$

where r_{pb} is the point-biserial correlation coefficient; P_1 represents the percentage of the dichotomous variable in the majority of cases; P_2 represents the percentage of the dichotomous variable in the minority of cases; and, y represents the "ordinate value (height of the curve on the y-axis)" which was located on the standard normal distribution table (Kline, 2005).

The independent variables examined in these analyses are located in the upper portion of the Activity System, Artifacts and Tools (see Figure 4.2).

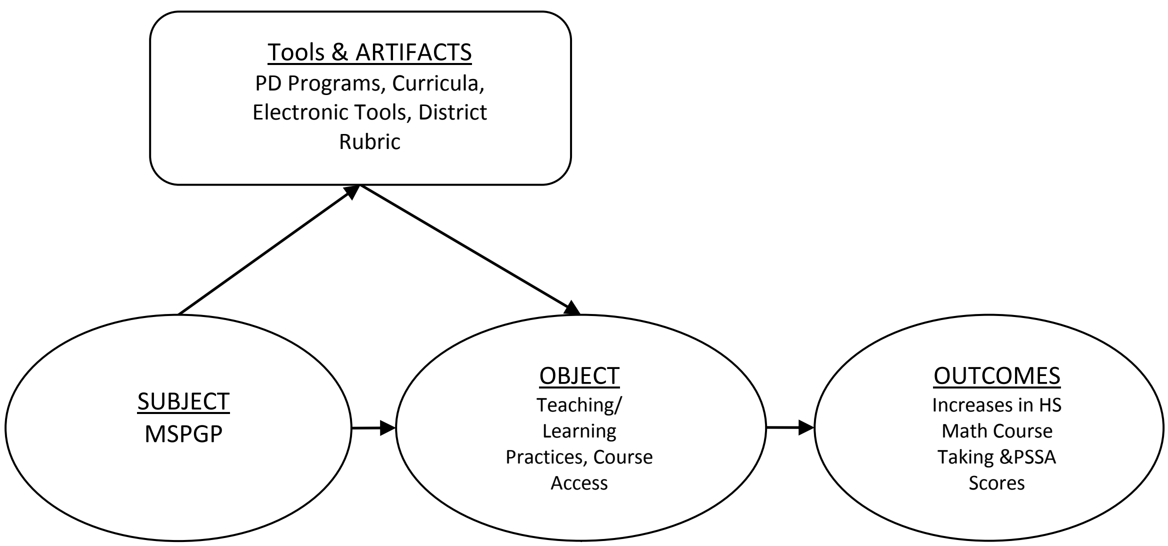


Figure 4.2. Activity Theory: Expanded Model: Upper and Central Regions

Results show that the relationship between the level of professional development and the overall increase in the percentage of students scoring at least proficient on the mathematics portion of the PSSA was significant, $r_p = .62$, $p < .05$. Moreover, the relationship between MSPGP engagement and the overall increase in the percentage of Latino/a students scoring at least proficient on the mathematics portion of the PSSA was also significant, $r_p = .90$, $p < .05$. All other PSSA correlations were positive, but weak to moderate (see Table 4.5). When changes in overall college prep mathematics course completion and higher level mathematics course completion were examined, all relationships appear to be weak. Values hovered around zero in both the negative and positive directions. When changes in achievement gaps for the PSSA were examined, results were both positive and negative and relatively weak. When changes in achievement gaps for the percentage of students completing college prep mathematics courses overall, as well as higher level, were examined, again, results were both positive and negative but weak.

Table 4.6. Biserial Correlations: MSPGP Participation and Changes in Student Outcomes

	Professional Development	MSPGP Engagement
	r_b	r_b
PSSA		
All Students	.62*	.40
African American Students	.36	.18
Latino/a Students	.50	.90*
CP Math Courses		
All Students	.30	.40
African American Students	-.08	-.11
Latino/a Students	-.02	-.03
Higher Level CP Math Courses		
All Students	.33	.31
African American Students	.08	.10
Latino/a Students	-.13	.13
Achievement Gaps		
PSSA		
White/African American Students	-.03	.30
White/Latino/a Students	-.34	-.26
CP Math Courses		
White/African American Students	.17	.33
White/Latino/a Students	.19	.01
Higher Level CP Math Courses		
White/African American Students	.06	-.10
White/Latino/a Students	.40	-.17

* $p < .05$ (2-tailed)

It appears that the relationship between participation in the MSPGP and changes in high school mathematics outcomes was a weak association, except for the change in PSSA outcomes. There was a significant moderate correlation for all students and a strong one for Latino/a students in particular. One result that is puzzling is that there appears to be no relationship between MSPGP participation and changes in PSSA scores for African American students, yet there was a significant increase in the percentage of African American students scoring at least proficient.

A serious limitation of the findings is that baseline data were not considered in the analyses. Recall that the MSPGP began working with schools and teachers during 2003. Although course completion data were collected during the 2003-04 academic year, much of the data were unreliable and therefore not considered in the analyses. Had this been a truly longitudinal examination, one that evaluated change over the five-year life of the grant, the results may have been different.

Correlations: Changes in Student Outcomes and District Stability

It may be the case that changes in outcomes were related to other factors which are located at the bottom of the Activity system (see Figure 4.3). In order to answer

Figure 4.3. Activity Theory: Expanded Model: Lower Region

question 2, correlations examined the relationship between some of the factors that are encompassed in: Rules and Norms, Community, and the Division of Labor, and the changes that occurred in outcomes. The factors that were investigated are those captured under the category of school district stability. They include: the type of curriculum; high school scheduling; the total minutes dedicated to teaching a full math course; and administrator turnover; including a change in the superintendent or principal.

The results revealed weak relationships among all variables, except for the relationship between the use of reformed curriculum and increases in PSSA scores, $r = .43$ $p = .043$ (see Table 4.7).

Table 4.7. Pearson Product Moment Correlations Between Changes in Student Outcomes and District Stability Variables

	Curriculum	Time	Schedule	Superintendent Turnover	Principal Turnover
	r	*r*	*r*	*r*	*r*
Changes in:					
PSSA	.43*	-.08	.35	-.07	-.11
CP Course Completion	-.14	.14	-.01	-.01	-.01
Adv. CP Course Completion	-.04	.11	.06	.11	-.17

* p =.043 (2-tailed)

Summary of Findings

Results showed that the number of schools that had an increase in the percentage of students scoring at least proficient was significant, $z = 2.5$, $p = .011$. Additionally, there was a significant increase in the percentage of students scoring at least proficient on the PSSA, $t = 4.03$, $df = 22$, $p = .001$. This was similar to the statewide trend in Pennsylvania. Moreover, the number of schools that had an increase in the percentage of African American students scoring at least proficient on the PSSA was significant, $z = 2.94$, $p = .003$ as well as the number of schools that had an increase in the percentage of

Latino/a students completing higher level mathematics courses, *z* = 2.91, *p* = .003. There was a significant increase in the percentage of African American students scoring at least proficient on the PSSA, *t* = 2.53, *df* = 13, *p* = .025. Again, this increase was similar to statewide the trend in Pennsylvania. There was also a significant increase in the percentage of Latino/a students completing higher level math courses *t* = 3.08, *df* = 19, *p* =.006. Biserial correlation results show that the relationship between the level of professional development and the overall increase in the percentage of students scoring at least proficient on the mathematics portion of the PSSA was significant, r_b= .62, $p < .05$ as was the relationship between MSPGP engagement and the overall increase in the percentage of Latino/a students scoring at least proficient on the mathematics portion of the PSSA, r_b= .90, $p < .05$. In examining the relationship between district stability variables and changes in outcomes, the results of the Pearson Correlation showed that there was a significant relationship between the mean overall increase in PSSA scores and the use of a reformed curriculum, *r* = .43, *p* = 0.43.

With regard to mathematics course completions, 70% of all schools had an increase in the percentage of students completing college prep mathematics courses and 65% had an increase in the percentage of students completing higher level mathematics courses. However, these increases were not significant, and there were no relationships between any of the increases in course completions and the amount of professional development a school participated in or the level of engagement schools had with the MSPGP.

The purpose of this chapter was to present the results of the analyses of a subset of MSPGP project data. The subset included 23 of the 46 school district partners in the MSPGP. First, descriptive data were presented to inform the reader about how and to what extent changes in student outcomes occurred from year 3 to year 5 of the project. Next, the biserial correlation was utilized to determine whether the level of engagement, high or low, that a district had in the project or the number of professional development hours, also high or low, math teachers participated in, had any relationship to those changes. Finally, other independent data, categorized as school district stability variables were examined to ascertain whether they had any relationship with changes in student outcomes. Correlations suggested that there were some relationships between changes in outcomes and the level of participation in the MSPGP; however, these relationships are not indicative of causation. Chapter Five will identify the limitations of this inquiry and make recommendations for further research on MSPs.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

Introduction

The Math Science Partnership program continues as a federal policy effort to remedy the chronic underachievement in mathematics and the sciences in the United States. The purpose of this study was to determine whether changes in high school mathematics outcomes had any relationship with participation in the Math Science Partnership of Greater Philadelphia. Activity Theory served as the lens through which the research was completed, providing both the theoretical and analytical perspectives for the study. There are several limitations to this investigation that will be discussed in this chapter. In addition, implications and recommendations will be discussed.

Discussion

An important finding of this study was that there was a significant increase in the percentage of all students scoring at least proficient on the PSSA. Moreover, there was a significant increase for African American students scoring at least proficient on the PSSA. These results were similar to statewide trends in Pennsylvania. There was also a significant increase for the percentage of Latino/a students completing higher level mathematics courses. However, when achievement gaps for African American and Latino/a students were examined, the results were disquieting with achievement gaps in the outcomes examined in this study increasing for several schools. These two, what appear to be dichotomous findings, open the door for an important discourse, a discourse captured in an article by Lubienski and Gutiérrez (2008) who suggest that in order to truly examine equity, both pieces of the puzzle are important: growth in achievement and opportunities to learn. Focusing solely on the significant gains that African American and Latino/a students had in MSPGP partner schools belies the fact that educational opportunity is not equitable across schools. It is important to look at the schools that experienced increases and determine what is in place in those schools that promoted such gains. Although beyond the scope of this study, this could be accomplished through case studies. "A greater emphasis on understanding contexts that produce excellence or gains in students who traditionally are not engaged by school may offer a different picture" (Lubienski & Gutiérrez p. 369) and may provide existence proofs for schools as they take on a commitment to equity and reform.

Davis and Martin (2008) believe that ultimately it is the teacher who has the most influence on students:

> Martin (2007) suggested [that teachers should] (a) develop a deep understanding of the social realities experienced by African American students, (b) take seriously one's role in helping to shape the racial academic, and mathematics identities of African American learners, (c) conceptualize mathematics not just as a school subject but as a means to empower African American students, and (d) become agents of change who challenge research and policy perspectives that construct African American children as less than ideal learners. (Davis & Martin, 2008, p. 23)

Reeves (2005) agrees that teachers play a vital role in student success. His research focused on poor and/or minority students from 228 different schools where 90% of the students were eligible for free or reduced lunch, 90% of the students were minorities, and 90% of the students "met or achieved high academic standards"(p. 186). He found that while there was no common formal program or intervention in place, teachers exhibited the following common behaviors: “A focus on academic achievement; clear curriculum choices; frequent assessment of student progress and multiple opportunities for improvement; an emphasis on nonfiction writing; and, collaborative scoring of student work” (p. 187).

The mission of the Math Science Partnership program includes reducing “achievement gaps in the mathematics and science performance of diverse student populations” (NSF, 2002, n. p.). To some extent, this occurred for several schools that were included in this study, but there is still much work that can be done through the MSP program. The literature suggests that culturally relevant pedagogy helps to improve mathematical understanding and problem-solving ability for students of color and English Language Learners (ELL) (Leonard, 2008; McClain & Berry, 2009).What was gleaned in this study with regard to improvement for the diverse student populations is important because it can set the groundwork to examine how and to what extent districts, schools, and teachers incorporate the cultural experience of African American and Latino/a students in their mathematics and science educational experience.

Another important finding is that there was a significant relationship between the level of professional development and the overall increase in the percentage of students scoring at least proficient on the mathematics portion of the PSSA. A correlation found a significant relationship between increases in PSSA scores and the use of reformed curriculum. Two different curricula, the Core-Plus Mathematics Project (CPMP) and the Interactive Mathematics Program (IMP) were implemented in 12 of the sample schools.

This finding contributes to the body of prior research on reformed curricula, which suggests that reformed curricula can provide positive learning experiences for students. According to Clark, Breed, and Fraser (2004):

> Students who have participated in the IMP program appear to be more confident than their peers in conventional classes; to subscribe to a view of mathematics as having arisen to meet the needs of society, rather than as a set of arbitrary rules; to value communication in mathematics learning more highly than students in conventional classes; and to be more likely than their conventionally-taught peers to see a mathematical element in everyday activity. These outcomes occurred while the IMP students maintained performance levels on the mathematics portion of the SAT at or above those of their peers in conventional classes. If student achievement outcomes are comparable, the mathematics education community must decide whether it values these consequences of a problem-based curriculum. (p. 7)

In Huntley, Rasmussen, Villarubi, Sangtong, and Fey's (2000) study of students from six different schools where the CPMP curriculum was field tested, they found that CPMP students when compared to control groups within the same schools, scored significantly higher on items that involved mathematics in real world contexts when using graphing calculators in Course 1 and Course 3 ($p < .001$, $p < .01$, respectively) and Course 2 without the use of calculators ($p < .001$). CPMP students also scored significantly higher on items that required model development and interpretation ($p < .001$). The control groups scored higher on items which were categorized as skill development and there were no differences on items that required students to evaluate expressions and solve equations and inequalities.

Webb (2003) found that IMP students scored significantly higher on items involving finite mathematics, probability and statistics than students in traditional math classes. He also found that students studying IMP took more math courses than did traditional students.

Boaler (2002) found that teachers of minority and working class students, who used reformed curricula “achieved a reduction in linguistic, ethnic and class inequalities in their schools” (p. 239) when they: helped students to understand the questions being asked of them; helped students to develop an appreciation for the importance of “written communication and justification” (p. 249); and “discussed with them ways of interpreting contextualized questions” (p. 250).

This brings to bear the importance of teaching practices in reformed classrooms. Teachers who focused on the relationship between the context of the mathematical learning situation and the experiences of students, and how those experiences influence the way in which students interpret mathematical problems were better able to judge the amount of support students needed in order to be successful. This meant giving explicit guidance in instances where students engage in reform-oriented practices that are not part of their repertoires of mathematical knowledge and skills, including providing opportunities for discussion about contextual information that may be unfamiliar to students. Moreover, when students’ reasoning skills were tied to their contextual lives, teachers provided the opportunities for students to explain both verbally and in writing how they went about reasoning. Finally, the teachers in this study did not completely rely on textbook problems because sometimes there was a contextual disconnect between the content of the textbook and the experience of the students. These teachers modified textbook problems and created problems that reflected the experiences of their students' lives.

Schoen, Cebulla, Finn, and Cos (2003) found that achievement on the "*Ability to Do Quantitative Thinking,* the mathematics subtest of the Iowa Test of Educational Development (ITED-Q)" increased their achievement "from the 52nd to the 56th percentile" (p.243) for students in CPMP classrooms when their teachers practiced reform-oriented pedagogy and when they held high expectations for all students. A regression model showed that the best indicator of achievement was whether or not the teacher participated in a two-week summer professional development focused on the curriculum.

While the findings presented here are positive, in most investigations which find improvement in achievement in classrooms using reformed curricula, teachers were supported in its implementation through extensive professional development. According to Remillard (2005): "Teachers require substantial support in learning to use new curriculum materials. They need to learn about the content, goals, approaches, and underlying assumptions of the curriculum they are being asked to use" (p. 239). In the MSPGP project, of the 11,273 hours of professional development provided to the teachers in this study, 5,653 hours were devoted to: curriculum alignment, curriculum development, or special training in curriculum. Although beyond the scope of this study, it would be interesting to delve deeper into the data and look for patterns of changes in student outcomes and the relationship they have to the type of professional development in which teachers participated. Furthermore, a comparison study could be completed to look for differences in achievement between students who were taught through a traditional curriculum and those who were taught through a reformed curriculum. Unfortunately, because the state of Pennsylvania has now mandated that students pass three end-of-course exams in order to graduate high school, the majority of schools in this study that used a reformed curriculum have since abandoned it, returning to the traditional course of study: Algebra 1, Geometry, Algebra 2, etc.

Limitations of the Study

The major limitation of this investigation is that individual student data were not collected by the project. Because of that fact, the unit of analysis was the school rather than the student. Thus, the sample size was 23, the number of Pennsylvania districts that were partners in the MSPGP for which the project had complete data sets, rather than 46,863, the approximate number of students who attended the 23 partner high schools. Moreover, the sample population of 23 schools represents only half of the 46 school districts that were MSGP partners and the study only investigated changes in mathematics outcomes on the high school level. MSPGP was involved in science education and also worked with middle schools. As such, this is a limited study of the partnership and therefore no valid conclusions can be drawn about the partnership as a whole.

The second limitation is that this evaluation captured only changes in outcomes form 2006 to 2008 the last three years of the project, rather than changes from baseline year 1 to year 5 because most of the data from year 1 were unreliable. In order to truly measure changes in outcomes that occurred during participation in the partnership, baseline data is needed.

A third limitation is in the analyses on changes in achievement gaps. Three Pennsylvania schools that were not included in the analyses comprised three of the five high schools in the partnership with the highest minority populations, at 54%, 61%, and 89%, respectively. None of the three could provide data on a yearly basis because they were severely understaffed and did not have the ability to devote the hours of time it took to assemble the data. This is an important finding in itself as it relates to equity and school funding. If school districts are expected to make data informed decisions about students, equitable funding is needed so that schools have the capacity to produce reliable data that can be acted upon judiciously.

The fourth limitation is that it was not possible to match the work of higher education mathematics faculty members to individual schools, beyond their time and effort spent in conducting professional development. Because one of the unique features of the MSP program is the participation of math and science disciplinary faculty in K-12 education, marrying IHE faculty work to individual schools would provide a finer grained analysis. In the annual reporting through the MSP Management Information System (MIS) system the project stated that: "We now wish we had developed a database for the collection of our IHE project-related activities" (WESTAT, 2008, n. p). Currently, a Co-Principal Investigator for the MSPGP is engaged in research on IHE involvement in MSPs.

The next limitation is in the interpretation of student outcomes. The outcomes investigated in the study were changes in a state mandated test and changes in the percentage of students completing mathematics courses. This is a very narrow definition of mathematics outcomes. While NCLB requires 100% proficiency in mathematics as measured by a standardized test, this cannot truly measure everything a student knows and is able to do. Likewise, a change in course completions tells little about what a student knows and is able to do. Having student level achievement data beyond percentages of course completions would provide for a much more thorough evaluation of the mathematical knowledge and skills that students developed and mastered while their schools were partners in the MSPGP.

The final limitation is the lack of qualitative data in the study. Without knowing what initiatives the schools were involved in besides being partners in the MSPGP, it is difficult to make a definitive case that involvement in the MSPGP had any relationship with changes in outcomes. Moreover, this study neglects the voice of the teacher. Interviews with teachers about the challenges and successes they experienced while being a member of the partnership would add depth to the story and could possibly point to other avenues for investigation. Furthermore, documentation of classroom observations would have helped to determine the level of the fidelity of implementation of the curricula and pedagogical practices in which teachers received professional development.

Educational Implications and Recommendations

The Math Science Partnership Program continues its mission throughout K-20 education and emergent research is increasingly important to sustain organizational capacity, and scaling up efforts. As the nation prepares for the reauthorization of the Elementary and Secondary Education Act, the MSP program is poised to assume an essential role in how math and science education progresses in the future. The purpose of this study was to contribute to the knowledge base that informs the MSP program and to explore the changes that occurred in high school mathematics outcomes for partner schools in the MSPGP.

One premise of the MSP is that when mathematicians, scientists and engineers from colleges and universities partner with K-12 institutions, implementing “change in mathematics and/or science educational practices in both higher education institutions and in schools and school districts,…[it will result] in improved student achievement across the K-12 continuum” (NSF, 2002, n. p.). Yet the extent to which this partnering is mutually beneficial is questionable. Pomeroy, Wolff and Rui (2010) analyzed Westat data for the entire MSP program that is generated yearly through a survey of higher education faculty and found that of the 605 faculty members who completed the survey, only 82 indicated that involvement in an MSP had a positive impact on the work they did with districts compared to 208 who indicated a positive impact on the work they did in teacher preparation and development. It is recommended that this be further investigated. What were the barriers that prevented the work at the district level from having a positive impact on faculty members? It is important to uncover this so that the MSP model is beneficial for all partners.

Two of the priorities in the Obama administration's *A Blueprint for Reform* (2010) are "Better Assessments" (US Department of Education, p. 3) and "Greater Equity" (p. 5). Therefore, a serious examination of how students are assessed is warranted. It is recommended that individual student level achievement data be collected in order to attain meaningful robust results. Moreover, individual student level data would allow MSPs to truly evaluate student growth over time. With student level data, it would be possible to determine early during a student's high school career whether or not he/she is on track for graduation on time and whether he/she is prepared for college. According to the *Blueprint*, the federal government is prepared to:

> support the development and use of a new generation of assessments that are aligned with college- and career-ready standards, to better determine whether students have acquired the skills they need for success. New assessment systems will better capture higher-order skills, provide more accurate measures of student growth, and better inform classroom instruction to respond to academic needs. (p. 3)

The MSP program has collected student data throughout the duration of the project, but the time has come for the data collection to be more thorough than it currently is. Individual student achievement data as well as changes in student achievement over time is warranted especially in this environment of increased accountability. Proficiency in mathematics will no longer only be a score on a standardized test. "State accountability systems will be asked to recognize progress and growth and reward success, rather than only identify failure" (US Department of Education, 2010, p. 9).

The *Blueprint* also recommends improvements in the professional development of teachers. It calls for states and districts to: "to support educators in improving their instructional practice through effective, ongoing, job-embedded, professional development that is targeted to student and school needs" (p. 15).

Professional development will most likely remain an integral part of the MSP program and it is recommended that the program adopt a Continuous Practice Improvement (CPI) model as a follow-up to its summer and yearlong institutes. This model is both ongoing and job-embedded. It emerged during the 1990s in the School District of Philadelphia (Schifter, 2008). In essence it was a teaching hospital model, where resident teachers received intense professional development in how to effectively incorporate the use of technology in their instructional practices and how to manage the use of computers in their classrooms. These resident teachers became experts whose classrooms were open to visiting teachers to learn about best practices. The teachers spent three days in the resident teacher's classroom and participated in a 30 hour program. Once they had completed this, they were further supported by a visiting technology specialist who provided pedagogical coaching, technical assistance, as well as ideas for lessons and activities.

Although teachers experienced some frustrations with the technology, overall they felt equipped to use it effectively in their classrooms. Moreover, the enthusiasm that the technology generated in their students' desire to learn was indeed a benefit. Although the CPI model ended in 2003, there are valuable lessons to be learned. Through CPI, communities of practice emerged within and between schools. This contributed to improving the school culture, thereby improving classroom culture as evidenced by the increase in enthusiasm of the students. "This is what CPI was all about – teachers teaching teachers, modeling best practices, all supported by principals who understood what it took to make change happen" (Schifter, p. 82). This model could help to build capacity within schools. This is important because participation in an MSP is only possible if the MSP continues to be funded.

One of the goals of the MSP program is to "Enhance the quality, quantity, and diversity of the K-12 mathematics and science teacher workforce" (NSF, 2003, n. p). Currently the MSPGP has extended its work through a Master Teacher Project. This project involves four teachers from six of the original 46 district partners. These teachers receive five graduate courses, three in leadership and two in content. They also receive a stipend in return for working in their schools to improve mathematics and science teaching and learning. Both models, CPI and the Master Teacher Project, can help increase quality by developing teacher leaders in MSP schools at the grass roots level so that there is on-going support from other teachers who share the same school culture.

Because schools are microcosms of society, each with its unique structure and culture, Activity Theory could aid in examining the systems and social networks that exist within districts and schools. These influences on policy and practice must be well understood at the outset of partnership formation. In order for MSPs to "significantly reduce achievement gaps in the mathematics and science performance of diverse student populations" (NSF, 2002, n. p.), a thorough examination of what makes a school what it is, is needed, otherwise, equity issues cannot be addressed. The Math Science Partnership of Greater Philadelphia partnered with 46 different school districts, each with its own unique culture and resources, or lack thereof. Although every effort was made to truly understand what the culture and needs of each district was, the sheer number of the school district partners made this somewhat prohibitive.

In conclusion, this study focused on a limited data set collected by the MSPGP. It is recommended that these data continue to be probed in order to look for other significant relationships as well as deficiencies that may exist in the MSP program. It is also recommended that case studies be completed on some of the schools that experienced significant gains in outcomes as well as on some that did not. Furthermore, the data should be examined to determine whether districts that had a very high level of engagement had greater increases in student outcomes than those who had a very low level of engagement. If the MSP program is to continue to be a federal effort to remedy the chronic underachievement in mathematics and the sciences in the United States, project development should incorporate some of the recommendations outlined in the *Blueprint for Reform* so that the MSP program ensures that it is keeping pace with the national agenda for mathematics and science education.

REFERENCES CITED

Abdal-Haqq, I. (1998). *Professional Development Schools: Weighing the Evidence.* Thousand Oaks, CA: Corwin.

Adelman, C. (2006). *The Toolbox Revisited: Paths to Degree Completion From High School Through College.* Washington, DC: U.S. Department of Education. Retrieved on September 22, 2009, from http://www.ed.gov/rschstat/research/pubs/toolboxrevisit/toolbox.pdf

Balfanz, R., & Byrnes, V. (2006). Closing the achievement gap in high-poverty middle schools: enablers and constraints. *Journal of Education for Students Placed at Risk, 11*(2), 143-159.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: W. H. Freeman.

Barab, S. A., Barnett, M., Yamagata-Lynch, L., Squire, K., & Keating, T. (2002). Using activity theory to understand the contradictions characterizing a technology-rich introductory astronomy course. *Mind, Culture, and Activity, 9*(2), 76-107.

Blank, R. K., & de las Alas, N. (2009). *Effects of Teacher Professional development on Gains in Student Achievement: How Meta Analysis Provides Scientific Evidence Useful to Education Leaders.* Council of Chief State School Officers (CCSO). Retrieved on November 13, 2009, from http://www.ccsso.org/content/pdfs/Final%20Meta%20Analysis%20Paper%20full.pdf

Boaler, J. (2002). Learning from teaching: exploring the relationship between reform curriculum and equity. *Journal for Research in Mathematics Education 33*(4), 239-258.

Book, C. L. (1996). Professional development in schools. In J. Sikula (Ed.), *Handbook on research on teacher education* (2nd ed., pp. 194-210). New York: Macmillan.

Brace, N., Kemp, R., & Snelgar, R. (2006). *SPSS for Psychologists, 3rd ed.* Hillsdale, NJ: L. Erlbaum Assoc.

Brazeel, P. L. (1981). Reagan cuts hit NSF, USDA, NIH. *Bioscience, 31*(4), 329-330.

Campoy, R. (2002). Building on expertise. In S. Mitchell (Ed.), *Effective Educational Partnerships: Experts, Advocates, and Scouts.* Westport, CT: Praeger Publications.

Carnegie Forum on Education and the Economy. (1986). *A Nation Prepared: Teachers for the 21st Century.* Washington, DC: The Task Force on Teaching as a Profession. (Eric Document Reproduction No. 268120)

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: an experimental study. *American Educational Research Journal, 26*(4), 499-531.

Castle, S., Arends, R. I, & Rockwood, K. D. (2008). Student learning in a professional development school and a control school. *The Professional Educator, 32*(1), n. p. retrieved on January 26, 2010, from http://www.theprofessionaleducator.org/articles/archives/spring2008.pdf

Castle, S., Fox, R. K., & Souder, K. O. (2006). Do professional development schools (PDSs) make a difference? a comparative study of PDS and non-PDS teacher candidates. *Journal of Teacher Education, 5*(1), 65-80.

Checchi, D. (2006). *The Economics of Education: Human Capital, Family Background, and Inequality*. New York: Cambridge University Press.

Choy, S. P., Chen, X., & Bugarin, R. (2006). *Teacher Professional Development in 1999–2000: What Teachers, Principals, and District Staff Report* (NCES 2006-305). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Clark, D., Breed, M., & Fraser, S. (2004). The consequences of a problem-based mathematics curriculum. *The Mathematics Educator. 14*(2), 7-16.

Clifford, M., & Millar, S. B. (with Smith, Z., Hora, M., & DeLima, L.). (2008). *K–20 partnerships: Literature review and recommendations for research* (WCER Working Paper No. 2008-3). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved September 23, 2009 from http://www.wcer.wisc.edu/publications/workingPapers/papers.php

College Entrance Examination Board (2009). 2009 profile of college-bound seniors. New York: Author. Retrieved on September, 200, 2009 from http://professionals.collegeboard.com/data-reports-research/sat/cb-seniors-2009

Commission on Teaching and America’s Future. (1996). What Matters Most: Teaching for America’s Future. New York: NCTAF. Retrieved on March 3, 2009, from http://www.nctaf.org/documents/WhatMattersMost.pdf

Committee on Science & Mathematics Teacher Preparation. (2001). *Educating Teachers of Science, Mathematics & Technology: Practices for the New Millennium*. Washington, DC: National Academy Press.

Cook, B. J. & Córdova, D. I. (2007). *Minorities in Higher Education Twenty-Second Annual Status Report: 2007 Supplement.* Washington, DC: American Council on Education. Retrieved on December 12, 2009, from http://www.acenet.edu/AM/Template.cfm?Section=CAREE&Template=/CM/ContentDisplay.cfm&ContentID=23716

Corcoran, T. B., Shields, P. M., & Zucker, A. A. (1998). *The SSIs and professional development for teachers*. Menlo Park, CA: SRI International. Retrieved on October 1, 2009, from http://policyweb.sri.com/cep/publications/ssiprfdv.pdf

Darling-Hammond, L. (1994). *Professional Development Schools: Schools for Developing a Profession*. New York: Teachers College Press.

Darling-Hammond, L., & Post, L. (2000). Inequality in teaching and schooling: supporting high- quality teaching and leadership in low-income schools. In R. D. Kahlenberg (Ed.), *A Notion at Risk: Preserving Public Education as an Engine for Social Mobility*. (pp. 127-167). New York: Century Foundation.

Davis, J. & Martin, D. B. (2008). Racism, assessment, and instructional practices: implications for mathematics teachers of African American Students. *Journal of Urban Mathematics Education. 1*(1), 10-34

Dimitrov, D. M. (2008). Initial trends in MSP-related changes in student achievement with MIS data. *Peabody Journal of Education, 83*(4), 637-653.

Doctoral Degree Awards to African Americans Reach Another All-Time High**. (**Winter, 2005/2006). *The Journal of Blacks in Higher Education*, *50*, 6-10.

Engestrom, R. (1987). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Helsinki: Orienta-Konsultit.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

Field, A. (2009). Discovering Statistics Using SPSS (3rd ed.). London: Sage.

Firestone, W. A, Mangin, M., & Martinez, M. C. (2007). Districts, teacher leaders, and distributed leadership: changing instructional practice. *Leadership & Policy in Schools*, *6*(1), 3-35.

Firestone, W. A, Mangin, M., Martinez, M. C., & Polovsky, T. (2005). Leading coherent professional development: a comparison of three districts. *Educational Administration Quarterly*, *41*(3), 413-448

Fisler, J. L., & Firestone, W. A. (2006).Teacher learning in a school–university partnerships: exploring the role of social trust and teaching efficacy beliefs. *Teachers College Record, 108*(6), 1155-1185.

Flores, A. (2007). Examining disparities in mathematics education: achievement gap or opportunity gap? *The High school Journal, 9*(1), 29-42.

Fullan, M. (2007). *The New Meaning of Educational Change* (4th Ed.). New York: Teachers College Press.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945.

Gill, B. & Hove, A. (2000). *The Benedum Collaborative Model of Teacher Education: A Preliminary Evaluation*. Santa Monica, CA: RAND. Retrieved on January 22, 2009, from http://www.rand.org/pubs/documented_briefings/DB303/

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context (NCES 2009–001 Revised). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved on September 20, 2009, from http://nces.ed.gov/pubs2009/2009001.pdf

Goodlad, J. I. (1991). Why we need a complete redesign of teacher education. *Educational Leadership, 49*(3), 4-10.

Goodlad, J. I. (1993). School-university partnerships and partner schools. *Educational Policy*, *7*(1), 24-39.

Guskey, T. R. (1997). Research needs to link professional development and student learning. *Journal of Staff Development 18*(2), retrieved on 11/30/07 from http://www.nsdc.org/library/publications/jsd/jsdgusk.cfm

Guskey T. R. (2002), Professional development & teacher change. *Teachers & Teaching: Theory & practice. 8*(3-4), 381-391.

Gutiérrez, R. (2008). A "gap-gazing" fetish in mathematics education? problematizing research on the achievement gap. *Journal for Research in Mathematics Education, 39*(4), 357-364.

Gutiérrez, R. (November, 2000). Advancing African-American, urban youth in mathematics: unpacking the success of one math department. *American Journal of Education*. *109*, 63-111.

Hamos, J. E. (2006). Framing K-12 partnerships in order to make a difference. *Academic Medicine, 81*(6), 11-14.

Heck, D. J., & Weiss, I. R. (2005). Strategic leadership for education reform: lessons from the statewide systemic initiatives program. *CPRE Policy Briefs.* (ERIC Document Reproduction Service No. ED389532) Retrieved on May 13, 2009, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/29/df/b8.pdf

Holmes Partnership. (2006). *Holmes Partnership Trilogy: Tomorrow's Teachers, Tomorrow's Schools, Tomorrow's Schools of Education*. New York: Peter Lang Publishing. Retrieved on August 17, 2009, from Education Research Complete.

Horn, I. S. (June, 2008). The inherent interdependence of teachers. *Phi Delta Kappan, 89*(10), 751-754.

Houston, W. R., Hollis, L. Y., Clay, D., Ligons, & C. M., Roff, L. (1999). Effects of collaboration on urban teacher education programs and professional development schools. In D. M. Byrd & D. J. McIntyre (Eds.), *Research on Professional Development Schools: Teacher Education Yearbook VI* (pp. 6-28) Thousand Oaks, CA: Corwin Press, Inc.

Humphreys, L. G., & Swets, J. A. (1991). Comparison of predictive validities measured with biserial correlations and ROCs of signal detection theory. *Journal of Applied Psychology*, *76*(2), 316-321.

Huntley, M. A., Rasmussen, C. L., Villarubi, R. S., Sangtong, J., Fey, J. T. (2004). Effects of standards-based mathematics education: a study of the Core-Plus Mathematics Project algebra and functions strand. *Journal for Research in Mathematics Education. 31*(3), 328-361.

Killion, J. (1998). Scaling the elusive summit. *Journal of Staff Development 19*(4), 12-15. Retrieved on September 28, 2009, from http://vnweb.hwwilsonweb.com.libproxy.temple.edu/hww/results/results_single_fulltext.jhtml

Kirschner, B. W., Dickinson, R., & Blosser, C. (1996). From cooperation to collaboration: the changing culture of a school/university partnership. T*heory Into Practice, 35*(3), 205-213.

Kline, T. J. B. (2005). *Psychological Testing: A Practical Approach to Design and Evaluation*. Thousand Oaks, CA: Sage Publication, Inc.

Knight, S. L., Wiseman, D. L., & Cooner, D. (2000). Using collaborative teacher research to determine the impact of professional development school activities on elementary students' math and writing outcomes. *Journal of Teacher Education, 51*(1), 26-38.

Koger, M. E., Thacker, A. A., & Dickinson, E. R. (April, 2004). Relationships Among the Pennsylvania System of School Assessment (PSSA) Scores, SAT Scores, and Self-Reported High School grades for the Classes of 2002 and 2003. Alexandria, VA: HumRRO. Retrieved on December 12, 2009, from http://www.portal.state.pa.us/portal/server.pt/community/technical_analysis/7447

Kramer, S. L., & Keller, R. (2008). An existence proof: successful joint implementation of the IMP curriculum and a 4 X 4 block schedule at a suburban U. S. high school. *Journal for Research in Mathematics Education, 39*(1), 2-8.

Kuutti, K. (1995). Activity theory as a potential framework for human-computer interaction research. In B. Nardi (Ed.) *Context and Consciousness: Activity Theory and Human Computer Interaction*, (pp. 17-44) Cambridge: MIT Press. Retrieved on September 20, 2008 from http://www.dwr.bth.se/kari_kuutti%20Nardi_book.pdf.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: understanding achievement in U.S. Schools, *Educational Researcher, 35*(7), 3-11.

Lee, J. (2002). Racial and ethnic achievement gap trends: reversing the progress towards equity? *Educational Researcher, 31*(1), 3-12.

Lee, J. (2004). Multiple facets of inequity in racial and ethnic achievement gaps. *Peabody Journal of Education, 79*(2), 51-57.

Leonard, J. (2008). *Culturally Specific Pedagogy in the Mathematics Classroom: Strategies for Teachers and Students.* New York: Routledge.

Leonard, J., Lovelace-Taylor, K., Sanford-DeShields, J., Spearman, P. (2004). Professional development schools revisited: reform, authentic partnerships, and new visions. *Urban Education, 39*(5), 561-583.

Leontj'ev, A. N. (1978). *Activity, Consciousness, and Personality*. Prentice-Hall.

Levin, B. B. & Rock, T. C. (2003). The effects of collaborative action research on preservice and experienced teacher partners in professional development schools. *Journal of Teacher Education, 54*(2), 135-149.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis. 15*(2), 129-151.

Luria, A. R. (1976). *Cognitive Development: Its Cultural and Social Foundations*. Cambridge, MA: Harvard University Press.

Liming D. & Wolf, M. (2008) Job outlook by education, 2006-16. *Occupational Outlook Quarterly,* Fall. Retrieved on September 20, 2009, from http://www.bls.gov/opub/ooq/2008/fall/art01.pdf

Lollor, C. (2000). Education: lessons about learning. *America's Investment in the Future*. National Science Foundation. Washington, DC Retrieved on April 20, 2009, from http://www.nsf.gov/about/history/nsf0050/index.jsp

Lomask, M. (1976). *A Minor Miracle: An Informal History of the National Science Foundation*. National Science Foundation (Report Number 76-18). Washington, DC Retrieved on April 20, 2009, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/39/e5/6b.pdf

Lubienski, S. T. (2008). On "*gap gazing*" in mathematics education: The need for gaps analyses. *Journal for Research in Mathematics Education, 39*(4), 350-356.

Lubienski, S. T. & Gutiérrez, R. (2008). Bridging the gaps in perspectives on equity in mathematics education. *Journal for Research in Mathematics Education*, *39*(4), 365-371.

Mazuzan, G. T. (1994). *The National Science Foundation: A Brief History*. (National Science Foundation Document nsf8816). Retrieved on April 20, 2009, from www.nsf.gov/about/history/nsf50/nsf8816.jsp

McClain, O. L., & Berry, R. Q. (2009). Examining the impact of divergent pedagogical styles on the mathematics learning of African American middle school students. In Swars, L. L., Stinson, D. W., & Lemons-Smith, S. (Eds.), Proceedings of the 31st annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Atlanta, GA: Georgia State University, Vol. 5 461-468.

McCormack, A. J. (1983). University/public school partnership: let's get on with it. *Biology Teacher*, *45*(8), 409.

McKeown, M., Klein, K., & Patterson, C. (2001). The National Science Foundation Systemic Initiatives. How a small amount of federal money promotes ill-designed mathematics and science programs in K-12 and undermines local control of education. In S. Stotsky, (Ed.), *What's at Stake in the K-12 Standards Wars: A primer for Educational Policy Makers*. (pp. 313-367). New York: Peter Lang. Retrieved on May 15, 2009, from http://libproxy.temple.edu:2280/ehost/pdf?vid=1&hid=102&sid=480ef8a9-b30f-48d2-8609-f1e15321f7ef%40sessionmgr109

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., et al. (1987). *The Underachieving Curriculum: Assessing U.S. School Mathematics From An International Perspective.* Champaign, Il: Stipes Publishing Company (ERIC Document Reproduction Service No. ED297930) Retrieved on May 13, 2009, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1c/c2/00.pdf

Mead, R., Smith, R.M., & Swanlund, A. (December, 2003). *Technical analysis: 2003 Pennsylvania System of School Assessment, Mathematics and Reading.* Data Recognition Corporation. NCLB. (2004). Retrieved on December 12, 2009, from http://www.portal.state.pa.us/portal/server.pt/community/technical_analysis/7447

Mervis, J. (1998). Mixed grades for NSF's bold reform of statewide education. *Science, 282*(5393), 1800-1805.

Mervis, J. (2001). Lawmakers vie to shape NSF program. *Science, 292*(5521), 1463 - 1465.

Moyer-Packenham, P. S., Bolyard, J. J., Kitsantas, A., & Oh, H. (2008). The assessment of mathematics and science teacher quality. *Peabody Journal of Education, 83*(4), 562-591.

Murphy, J. (2009). Closing achievement gaps: lessons from the last 15 years. *Phi Delta Kappan, 91*(3), 8-12.

National Center for Education Statistics (2009).The Nation's Report Card: Mathematics 2009 (NCES 2010–451). Institute of Education Sciences, U.S. Department of Education, Washington, D.C. Retrieved on October 15, 2010, from http://nces.ed.gov/nationsreportcard/pdf/main2009/2010451.pdf

National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform.* United States Department of Education. Retrieved on April 28, 2009, from http://www.ed.gov/pubs/NatAtRisk/risk.html

National Commission on Teaching and America's Future (NCTAF). (1996). *What Matters Most: Teaching for America's Future*. New York: National Commission on Teaching and America's Future. Retrieved on August 18, 2009, from http://www.nctaf.org/documents/WhatMattersMost.pdf

National Council for Accreditation of Teacher Education (NCATE). (2001). *Standards for Professional Development Schools*. Washington, DC: Author. Retrieved on September 8, 2009, from http://www.ncate.org/public/pdsOverview.asp?ch=136

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics.* Reston, VA: Author.

National Science Board. (2006). *Science and Engineering Indicators 2006.* Two volumes. Arlington, VA: National Science Foundation (volume 1, NSB 06-01; volume 2, NSB 06-01A). Retrieved on February 14, 2010, from http://www.nsf.gov/statistics/seind06/pdfstart.htm

National Science Foundation. (2002). *Math and Science Partnership Program (MSP) Program.* (Program solicitation, NSF-02-061). Arlington, VA: Author. Retrieved on February 22, 2009 from, http://www.nsf.gov/pubs/2002/nsf02061/nsf02061.html

National Science Foundation. (2003). *Math and Science Partnership Program (MSP) Program* (Program solicitation, NSF-03-605). Arlington, VA: Author. Retrieved on February 22, 2009 from, http://www.nsf.gov/pubs/2003/nsf03605/nsf03605.htm

National Science Foundation. (2004). *National Science Foundation Strategic Plan FY 2003-2008.* (NSF004-201) Arlington, VA: Author. Retrieved on February 22, 2009, from http://www.nsf.gov/pubs/2004/nsf04201/FY2003-2008.pdf

National Science Foundation. (2006). *Math and Science Partnership (MSP).* (Program solicitation, NSF-06-539). Arlington, VA: Author. Retrieved on June 7, 2009, from http://www.nsf.gov/pubs/2006/nsf06539/nsf06539.pdf

National Science Foundation (2009). *Math Science Partnership (MSP)* (Program Solicitation, NSF-09-507). Arlington, VA: Author. Retrieved on July 19, 2009, from http://www.nsf.gov/pubs/2009/nsf09507/nsf09507.pdf

Nettles, M. T. & Millett, C. M. (1999). *The Human Capital Liabilities of Underrepresented Minorities in Pursuit of Science, Mathematics, and Engineering Doctoral Degrees.* Chicago, IL: Spencer Foundation. (ERIC Document Reproduction No. 440576).

Neuman, W. L. (2002). *Social Research Methods: Qualitative and Quantitative Approaches* (5th Ed.). Upper Saddle River, NJ: Pearson.

No Child Left Behind Act Public Law PL 107-110. (2001). U. S. Department of Education. Retrieved on October 5, 2008 from, http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Noffke, S. E., Clark, B. G., Palmeri-Santiago, J., Sadler, J., & Shujaa, M. (1996). Conflict, learning, and change in a school/university partnership: different worlds of sharing. *Theory Into Practice, 35*(3), 165-172.

Osguthorpe, R. T. & Patterson, R. S. (1998). *Balancing the Tensions of Change.* Thousand Oaks, CA: Corwin Press, Inc.

Ou, D. (2009). To leave or not to leave? A regression discontinuity analysis of the impact of failing the high school exit exam. *Economics of Education Review*. *6*(2), 1-16.

Pane, J. F., Williams, V. L., Olmsted, S. S., Yuan, K., Spindler, E., & Slaughter, M. E. (2009). Math Science Partnership of Southwest Pennsylvania: Measuring Progress Toward Goals. Santa Monica, CA: RAND. Retrieved on July 22, 2009, from http://www.rand.org/pubs/monographs/2009/RAND_MG857.pdf

Pennsylvania Department of Education (2009). *PSSA and AYP Results 2008-2009.* Retrieved on February 27, 2010, from http://www.state.pa.us/portal/server.pt/community/school_assessments/7442/2008-2009_pssa_and_ayp_results/600286

Pine, G. J. (2003). Making a difference: a professional development school's impact on student learning. In D. Weisman, S. Knight (Eds.), *Linking School-University Collaboration and K-12 Student Outcomes.* (pp. 31-47) N.Y: AACTE Publications.

Pomeroy, D, Wolff, E. F., & Rui, N. (2010) *Impact of math science partnership work of higher education faculty*. Paper presented at the MSP Learning Network Conference, Washington DC. Retrieved on August 1, 2010 from http://hub.mspnet.org/media/data/37_Pomeroy.pdf?media_000000006865.pdf

Rasch, K. D. (1999). Ripples and waves: the growth of Parkway South High School-Maryville University Partnership. *Peabody Journal of Education, 74*(3&4), 71-84.

Remillard, J. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research. 76*(2), 211-246.

Reeves, D. B. (2005). In *High performance in high poverty schools: 90/90/90 and beyond. Accountability in Action,* 2nd ed. pp. 185-208. Englewood, CO: Advanced Learning Press.

Ridley, D. S., Hurwitz, S., Hackett, M. R. D., & Miller, K. K. (2005). Comparing PDS and campus-based preservice teacher preparation: is PDS-based preparation really better? J*ournal of Teacher Education, 56*(1), 46-56.

Rigden, D. W. (1991). *Business /School Partnerships: A path to Effective Restructuring*. (2nd ed.). New York: Council for Aid to Education.

Rose, H. & Betts, J. R. (2004). The effect of high school courses on earnings. *The Review of Economics and Statistics*, *86*(2), 497-513.

Ross, J. A., Bruce, C., & Hogaboam-Gray. (2006). The impact of a professional development program on student achievement in grade 6 mathematics. *Journal of Mathematics Teacher Education, 9(6),* 551-577.

Russel, D.R. (2002). Looking beyond the interface: activity theory and distributed learning. In: M.R. Lea & K. Nicoll (Eds.). *Distributed Learning. Social and Cultural Approaches to Practice*. Routledge Falmer.

Sanders, J. (2005). An activity theory perspective. *Work Based Learning in Primary Care*, *3*, 191-201.

Sayler, B., & Apaza, J. (2006). *Using Multiple Measures to Guide Mathematics Reform Within a K-12 District*. Center for the Advancement of Mathematics and Science Education. Spearfish, SD: Black Hills State University. Retrieved on May 18, 2009, from http://prime.mspnet.org/media/data/SaylerApaza.pdf?media_000000006113.pdf

Saxe, G. B., Gearhart, M., & Nasir, N. S. (2001). Enhancing students' understanding of mathematics: a study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education, 4*(1), 55-79.

Scherer, J. (2008). A review of instruments to evaluate partnerships in math and science education. *Peabody Journal of Education*, *83*(4), 611-636.

Schifter, C. C. (2008). *Infusing Technology into the Classroom: Continuous Practice Improvement.* Hershey, PA: Information Science Publishing,

Seaman, J. (2008). Adopting a grounded theory approach to cultural-historical research: conflicting methodologies or complementary methods? *International Journal of Qualitative Methods*, *7*(1), 1-17.

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Cos, F. (2003). Teacher variables that relate to student achievement when using a standards-based curriculum. *Journal for Research in Mathematics Education. 34*(3), 228-259.

Silverstein, G., Bell, R., Frechtling, J., & Miyaoka, A. (2005). *MSP MIS Summary Data for Comprehensive and Targeted Partnership Projects: 2002–03 and 2003–04 School Years*. Rockville, Maryland: WESTAT. Retrieved on May 13, from, http://focus.mspnet.org/media/data/MIS_MSP_Report_Aug05.pdf?media_000000006086.pdf

Snow-Renner, R. & Lauer, P. A. (2005). Professional Development Analysis. Denver, CO: McREL. Retrieved on August 31, 2010, from www.mcrel.org./ProfessionalDevelopment/5051IR_Prof_dvlpmt_analysis.pdf

Stillwell, R. (2009). *Public School Graduates and Dropouts From the Common Core of Data: School Year 2006–07* (NCES 2010-313). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved February 14, 2010, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010313

Struchens, M. E. (2000). Confronting beliefs and stereotypes that impede the mathematical empowerment of African American students. In M. E. Struchens, M. Johnson, & W. Tate (Eds.) *Changing the Faces of Mathematics: Perspectives on African Americans* (pp. 7-14). Reston, VA: National Council of Teachers of Mathematics.

Swars, S. L., Meyers, B., Mays, L. C., & Lack, B. (2009). A two-dimensional model of teacher retention and mobility: classroom teachers and their university partners take a clo9ser look at a vexing problem. *Journal of Teacher Education, 60(2),* 168-183.

Teitel, L. (2004). *How Professional Development Schools Make a Difference: A Review of Research.* (2nd ed.). Washington, DC: NCATE.

Tushnet, N. C. (1993). *A Guide to Developing Educational Partnerships*. Office of Educational Research and Improvement, (ED). Washington, DC Los Alamitos, CA: Southwest Regional Lab. (ERIC Document Reproduction Service No. ED362992)

Tyler, J. H., & Vitanova, S. (2008). Does MSP participation increase the supply of math teachers? Developing and testing an analytical model. *Peabody Journal of Education, 83*(4), 536-561.

U.S. Department of Education, National Center for Education Statistics. (2009). *Digest of Education Statistics, 2008* (NCES 2009-020).

U.S. Department of Education (2010). *A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act.* (2010). Washington, DC. Author. Retrieved on September 11, 2010, from http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

University of Helsinki, Center for Activity Theory and Developmental Work Research. *Introduction to the Activity Theory.* Retrieved May 25, 2008, from http://www.edu.helsinki.fi/activity/pages/chatanddwr/activitysystem

Valdez, J. R. & Snyder, J. (2006). Social, cultural, and political influences on the development of an educational partnership. *Journal of Latinos and Education, 5*(1), 29-47.

Vanneman, A., Hamilton, L, Anderson, J. B., & Rahman, T. (2009). *Achievement Gaps: How Black and White students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress*. (NCES 2009-455). Retrieved on December 15, 2009, from http://nces.ed.gov/nationsreportcard/pubs/studies/2009455.asp

Villaseñor, A. & Kepner, H. S. (1993). Arithmetic from a problem-solving perspective: an urban implementation. *Journal for Research in Mathematics Education, 24*(1), 62-69.

Voorhees, R., & Lee, J. (2009). *Basis of Longitudinal Cohort Analysis: Principles and Practices of Student Success.* Indianapolis, ID: Lumina Foundation. Retrieved on March 8, 2010, from http://www.achievingthedream.org/docs/guides/ATD_Longitudinal-Cohort-Analysis.pdf

Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wailes, J. R. (1968). *History and Development of National Science Foundation Elementary Institutes.* U.S Department of Health, Education, and Welfare, Office of Education. (ERIC Document Reproduction Service No. ED027230)

Wang, J., & Goldschmidt, P. (2003). Importance of middle school mathematics on high school students' mathematics achievement. *The Journal of Educational Research, 97*(1), 3-13.

Waterman, A. T. (1960). National science foundation: a ten year resume. *Science, 131*(3410), 1342-1354.

Web, N. L. (2003). The impact of the Interactive Mathematics Program on student learning. In S. L. Senk & D. R. Thompson (Eds.) *Standards-based school mathematics curricula: What are they? What do students learn?* (pp. 375-398. Mahwah, NJ: Lawrence Erlbaum Associates.

Weber, M. (2003). *The History of Commercial Partnerships in the Middle Ages*. (L. Kaelber, Trans.) MD: Rowman & Littlefield. (Original Work: The Protestant Ethic and the Spirit of Capitalism, published 1905).

WESTAT. (2008). Math Science Partnership Management Information System. [Computer Software]. Washington, DC. Retrieved on September 15, 2010, from www.msp-mis.org.

Wilson, S. (2003). *California Dreaming: Reforming Mathematics Education*. New Haven/London: Yale University.

Wolfle, D. (1957). National science foundation: the first six years. *Science 126*(3269), 325-343.

Wong, K. K., & Socha, T. (2008). A pilot study to identify comparison schools for math and science partnership participating schools. *Peabody Journal of Education, 83(4)*, 654-673.

Yin, R. K., Hackett, E. J., & Chubin, D. E. (2008). Discovering "what's innovative": the challenge of evaluating education research and development efforts. *Peabody Journal of Education*, *83*(4), 674-690.

Yoon, K. S., Duncan, T., Lee, S. W.Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved on October 1, 2009, from http://ies.ed.gov/ncee/edlabs

Zucker, A. A., Shields, P. M., Adelman, N. E., Corcoran, T. B., Goertz, M. E. (1998). *A Report on the Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Programs. NSF Contract No. SED-9255371*. Retrieved on May 16, 2009, from http://www.nsf.gov/pubs/1998/nsf98147/nsf98147.pdf

APPENDIX A
SAMPLE MATH COURSE COMPLETION TABLE

Table 1a Mathematics Passing Rates by Level and Grade (All Students)

High School 2007-08	Grade 9		Grade 10		Grade 11		Grade 12		All Grades	
	N	%	N	%	N	%	N	%	N	%
Level 1	294	62%	16	3%	0	0%	1	0%	311	16%
Level 2	124	26%	323	61%	21	4%	7	1%	475	24%
Level 3	1	0%	148	28%	137	27%	20	4%	306	15%
Level 4	0	0%	0	0%	141	28%	45	9%	186	9%
Level 5	0	0%	0	0%	46	9%	27	6%	73	4%
AP Statistics	0	0%	0	0%	0	0%	8	2%	8	0%
AB AP Calculus	0	0%	0	0%	15	3%	7	1%	22	1%
BC AP Calculus	0	0%	0	0%	0	0%	15	3%	15	1%
Intro. to Calculus	0	0%	0	0%	84	16%	30	6%	114	6%
Statistics	0	0%	0	0%	1	0%	64	13%	65	3%
Probability & Statistics	0	0%	0	0%	3	1%	84	17%	87	4%
All College Prep Math	**419**	**88%**	**487**	**92%**	**448**	**88%**	**308**	**63%**	**1,662**	**83%**
Total Students	**477**		**528**		**512**		**487**		**2,004**	

* Data Source: MSPGP Data Collection Survey

APPENDIX B
SCHOOL DISTRICT RUBRIC

The Math Science Partnership of Greater Philadelphia

A Program of the National Science Foundation at La Salle University

Summary Report

Regional Coordinator:
Date Completed:
District:

Middle School Math	**Middle School Science**
Status: Struggling ☐ Status Quo ☐ Getting Ready ☐ On the Road ☐ Already Reformed ☐ **Level of Engagement:** None ☐ Low ☐ Moderate ☐High ☐ Very High ☐ **Summary of Evidence:**	**Status:** Struggling ☐ Status Quo ☐ Getting Ready ☐ On the Road ☐ Already Reformed ☐ **Level of Engagement:** None ☐ Low ☐ Moderate ☐ High ☐ Very High ☐ **Summary of Evidence:**
High School Math	**High School Science**
Status: Struggling ☐ Status Quo ☐ Getting Ready ☐ On the Road ☐ Already Reformed ☐ **Level of Engagement:** None ☐ Low ☐ Moderate ☐High ☐ Very High ☐ **Summary of Evidence:**	**Status:** Struggling ☐ Status Quo ☐ Getting Ready ☐ On the Road ☐ Already Reformed ☐ **Level of Engagement:** None ☐ Low ☐ Moderate ☐ High ☐ Very High ☐ **Summary of Evidence:**

MS Math ☐ HS Math ☐ MS Sci ☐ HS Sci ☐	Already Reformed	On the Road to Reform	Getting Ready	Status Quo	Struggling
Training	☐ Already received 130-240 hrs training in exemplary curriculum & student centered pedagogy	☐ Still receiving basic training (130-240 hrs) in an exemplary curriculum and student centered pedagogy	☐ May experiment with trial replacement units and limited PD of staff	☐ PD sessions are regular, but may lack coherence or focus.	☐ District focus is on coping strategies. PD sessions are disjointed or isolated
Classroom Practice	☐ Training is reflected in classroom practice	☐ Training is not always reflected in classroom practice	☐ Isolated teachers use constructivist practice. Many teachers think they are using such practices, but fall short of goals	☐ Dominated by recitation or other traditional practice. Very little differentiation of instruction.	☐ Teachers practice is focused on classroom management and basic skills
Cross-grade articulation	☐ Articulation has been achieved	☐ Articulation has not been achieved but is being worked on	☐ Articulation has been raised as an issue and adopted as a goal by majority of staff. Little or no actual articulation has occurred	☐ Limited to non-existent. Teachers feel uncomfortable outside of own specialty	☐ No cross-grade conversation at all, teachers are isolated individuals
Daily lesson pacing	☐ Pacing of lessons has been fine-tunes to yield crisp, optimally-paced, variegated lessons	☐ Not yet fine tuned to yield crisp, optimally-paced, variegated lessons	☐ Lesson timing is often awkward as teachers experiment with new activities or techniques	☐ Lesson pacing is entirely determined by the teacher. Activities may be rushed as a result of teacher talk	☐ Lessons frequently interrupted. Lessons slow for remediation and then rush to finish unit
Level of student challenge	☐ All students being challenged to their fullest capacity. No inequity in course taking by race, gender, or SES	☐ Most students are being challenged to their fullest capacity. Inequity in course taking by race, gender, or SES reduced to less than 1 year	☐ Steps taken to reduce inequities in level and amount of course taking, but only limited progress.	☐ Only those students who are exactly on pace with the lessons are being adequately challenged. Others are either bored or lost. Inequity >1year in level and amount of student course taking by race, gender, or SES	☐ Very few students are being challenged. Significant inequity (>1.5 years) exists in level and amount of student course taking by race, gender, or SES
Alignment of assessment with state standards & tests	☐ Assessment and curriculum aligned with state standards and tests	☐ Assessment and curriculum are still being aligned with state standards and tests	☐ Have not yet committed to changing curriculum, but are beginning to engage various stakeholders in discussions	☐ Curriculum may have been aligned in content domains, but in a shallow way and without significant alignment in pedagogical domains (e.g. inquiry,	☐ Teachers unfamiliar with Standards or assessment frameworks (such as Anchors in PA)

MS Math ☐ HS Math ☐ MS Sci ☐ HS Sci ☐	Already Reformed	On the Road to Reform	Getting Ready	Status Quo	Struggling
				modeling)	
Data collection	☐ Data on pre and post student achievement has been received & validates or invalidates cost benefit of reform efforts	☐ Data on student achievement pre and post reform or comparisons of reform and non-reform classes is in the process of being collected	☐ As a result of low scores on high stakes tests or a self analysis, district realizes their students are underachieving in certain areas.	☐ Data collection and analysis is focused on gap analysis rather than on systemic improvement	☐ Data collection is focused solely on AYP
Hiring, induction & evaluation practices	☐ Hiring, induction and evaluation practices are all aligned with reform efforts	☐ New hiring, induction and evaluation practiced have been established and aligned with reform efforts, but implementation is problematic	☐ New hiring, induction, and evaluation models are being developed or piloted, but are not established district-wide	☐ New hiring and induction does not consider preparation to teach constructivist lessons. Evaluation is based on traditional norms of classroom management	☐ District has large numbers of new hires every year to replace high staff turnover. Induction is focused on policy and behavior-management
Attitude of teachers	☐ Supportive. Teachers view themselves as advanced learners of content and pedagogy	☐ Not all of those "honestly skeptical" have been persuaded. Some teachers still resistant	☐ Some teacher leaders who want to make some kind of positive change in their curriculum & instruction methods. Other pockets of resistance.	☐ Teachers are either satisfied with their students' performance or blame students' poor performance on external factors beyond their control	☐ Teachers believe they are in a crisis. They may also be in the throes of labor unrest. Districts in "empowerment" could be in the midst of a reconstitution under private management or under state takeover.
Support from administration	☐ Superintendent to principals all understand and support the reform	☐ Some administrators are supportive and some are not. Some are resistant	☐ Some administrators want to make some kind of positive change in their curriculum & instruction methods	☐ Administrators cede curricular decisions to teachers as long as AYP is maintained	☐ Schools want to improve but have multi-system problems e.g. lack essential math and science supervisory staff or suffer from great instability in their administrative and/or teaching staffs.
Perspective of parents & school board	☐ Parents and school board members are pleased	☐ Parents and school board members are pleased for the most part, but there are still some quarters of discontent	☐ Parents and school board members may be resistant as district gets more serious about thinking	☐ Parents of successful students are well-satisfied. Parents of unsuccessful students are	☐ Parents are disenfranchised. School Board members resist changes not directly

MS Math ☐ HS Math ☐ MS Sci ☐ HS Sci ☐	Already Reformed	On the Road to Reform	Getting Ready	Status Quo	Struggling
			about change	likely to blame the student rather than the curriculum	connected with higher test scores in their own mental models
Student attitudes	☐ Students are enjoying learning, desire more math and or science and are keeping their options open about post-secondary studies & careers in math & science related fields	☐ Most students enjoy learning, desire more math and/or science and are keeping their options open. Some worry if they are being adequately prepared. Some blame their failure or lower grades on the new curriculum	☐ Some students may be upset by changes, especially those who were successful before	☐ Student motivation varies with history of success, but most students opt not to challenge themselves beyond what is required.	☐ Students are disaffected and disenfranchised. Apathy and antagonism are rampant

APPENDIX C
SCHOOL REFORM STATUS RUBRIC DESCRIPTIONS

"Already Reformed": This is a school whose math or science teaching staffs have already received extensive basic training (130-240 hours) in an exemplary curriculum and student centered pedagogy. Such training is reflected in actual classroom practice. Cross grade articulation has been achieved and daily lesson pacing has been fine-tuned to yield crisp, optimally-paced, variegated lessons. As a result, all students are being challenged to their fullest capacity. Assessment and curriculum units have been aligned with state standards and the state tests. Data on pre and post student achievement has been collected that either validates or invalidates cost benefits of the reform efforts. New teaching hiring, teacher induction and teacher evaluation practices have been established and aligned with the reform effort. Administrators from the superintendent to the building principals all understand and support the reform. Parents and school board members are pleased. Student are enjoying learning, desire more math and/or science and are keeping their options open about post secondary studies and careers in math and science related fields.

New Opportunities: Teachers are ready to be trained as reform leaders within and beyond their schools. The school becomes a model for other schools. Enhanced teacher professional development in more advanced math and science content, applications and technology are now possible. Math, science lesson integration, along with other disciplines, is a reachable goal within a school and district. Mutual, two-way partnerships with other entities can fully mature.

Challenges: New teacher and administrator turnover threatens to erode program gains if the same degree of original baseline training is necessary. Curriculum and instruction need to continually adapt to new social goals, better prepared students and state tests. Teachers who stop professional development can slide back into teaching tired and stale lessons. As a result, over time the school culture regresses, ossifies and reverts to what is most expedient rather than what optimizes student learning.

"On the Road to Reform": This is a school whose math and/or science teachers are still receiving basic training (130-240 hours) in an exemplary curriculum and student centered pedagogy. The training and use of the curriculum is not always reflected in actual classroom practice. Cross grade articulation has not been achieved but it still being worked on. Daily lesson pacing has not been fine-tuned to yield crisp, optimally-paced, variegated lessons. Classroom management can still be an issue is some classes. As result, not all students are being challenged to their fullest capacity. Assessment and curriculum units are still being aligned with state standards and the state tests. Data on student achievement pre and post reform or comparisons of reform and non-reform classes is in the process of being collected. Not all of those "honestly skeptical" have been persuaded. Some teachers and administrators are still resistant to reform. New teaching hiring, teacher induction and teacher evaluation practices have been established and aligned with the reform effort but implementation is problematic. Some administrators from the superintendent to the building principals understand and support the reform effort some do not. Parents and school board members are pleased for the most part but there may still be some quarters of discontent. Most student are enjoying learning, desire more math and/or science and are keeping their options open about post secondary studies and careers in math and science related fields, but some students still worry if they are being adequately prepared. Some others blame their failure or lower grades on the new curriculum.

New Opportunities: School test scores should start to show improvement and in some cases dramatic gains. New teaching and administrative leaders begin to emerge. Teachers grow as professionals. A new culture of data based decision making begins to take hold as those people who are honestly skeptical demand objective data to validate the reform effort. The school is devoting large blocks of time to substantive curricular based professional development. Principals are being stimulated to become more instructional leaders than mere building managers. Parents and school board members are undergoing some fundamental shift in thinking about math and science learning.

Challenges: New teacher and administrator turnover threatens to erode program gains if the same degree of original basic training is necessary. Skepticism and in some cases stubborn resistance to change are political obstacles. Purposeful lesson organization, curriculum pacing and student centered instructional technique needs to improve to realize optimal student learning opportunities. School and departmental culture need to solidify as a new learning community.

“Getting Ready”: These are schools that have some key administrator and teacher leaders who want to make some kind of positive change in their curriculum and instruction methods. The school could have low scores on high stakes tests or a district could have done a self analysis and come to the conclusion their students were underachieving in certain key areas. These administrators and teacher leaders are open to change but still researching and exploring options. They may do trial replacement units of new curricula and limited professional development of their staff. They have as yet no definite commitment to change their curriculum or nor are they ready to provide for extensive professional development but are getting ready to change by engaging their various stakeholders in the process. They may encounter staff resistance to change from some quarters or the resistance may emerge only if the staff and parents see that the school is serious about the change, or when it directly affects them. Typically, schools in this stage need a great deal of guidance and assistance from their “critical friends” who as partners help guide the school through the “getting ready” stage to reach a decision and made a tangible commitment to change.

Opportunities: Schools have the opportunities to make changes in curriculum, instruction assessment and school culture that could lead to dramatics gains in student achievement and attitudes toward continued math and science learning.

Challenges: The key hurdle is to come to as decision to pursue a specific course of action in math and science reform over multiple years and to commit the time and effort and resources to stay the course. Achieving a consensus about this course of action is very difficult but the “weight” needs to be clearly behind the reform or it will be stillborn.

“Status Quo” Schools that are status quo either see no need to change any of their practices because their students are already fully challenged and are scoring well, or if their students are not performing optimally, they do not want to spend the time and resources necessary to change.

Opportunities: As more schools become “Fully Reformed” and the regional culture of student and parents change thorough the region where higher levels of math and science achievement and understanding becomes more pervasion, status quo school will come under increased pressure to change. As teachers and administrators turnover in these schools and new personnel are hired who are familiar with and support reform, status quo schools may realize they need to change as well.

Challenges: Status quo schools can be a negative influence, either passively or actively, by providing resistance support to certain elements within other schools that are attempting math and science reform. For example, if only low achieving urban schools with disadvantaged students are attempting reform, while all the higher achieving schools with more socio-economically advantaged students in suburban areas remain status quo schools, the default model for success will be the status quo schools. It is therefore necessary to have a heterogeneous mix of schools engaged in reform to serve as “existence proofs” that all students can be challenged more even “honors students” and that “status quo” is not the model.

“Struggling”: These are schools that want to improve but have major and multi system problems. For example, schools and districts could lack essential math and science supervisory staff, or suffer from great instability in their administrative and/or teaching staffs. They could be in a financial crisis, in the throes of labor unrest, or they could be in the midst of a reconstitution under private management or a state takeover.

Opportunities: Struggling schools are often those with the neediest students who score among the lowest on state tests. Thus, there is large upside potential to significantly narrow inequities suffered by large number of students at a time. By having a heterogeneous group of schools at different places “on the road to reform”, struggling schools can have hope that curriculum and instruction can make substantially improved to increase student understanding and performance in math and science.

Challenges: The challenges facing struggling schools are all those previously listed, although more so. Paradoxically, low test scores in many struggling schools do not translate into widespread receptivity for change in curriculum and instruction. We have found individuals with the highest resistance to change are often located in struggling schools and districts. Few are willing to own up the failure, so the locus of responsibility is usually placed outside the classroom, school or district in the case of school board members. Thus, the most difficult challenge in struggling schools is to encourage, persuade and chide them to “own” the problem. This requires a great deal of facilitated dialogue that can often resemble family therapy session.

APPENDIX D
TEACHER PROFESSIONAL DEVELOPMENT HOURS

School District	*Academic Year*	*Total Number of Eligible Teachers*	*Total Number of Participating Teachers*	*Percent of Eligible Teachers who Participated*	*Total Professional Development Hours*	*Average Professional Development Hours*
Anthony City	2003-04	49	0	0%	0	0
	2004-05	47	28	60%	452.5	16.16
	2005-06	48	18	38%	200.5	11.14
	2006-07	47	34	72%	328.5	9.66
	2007-08	45	2	4%	35.5	17.75
	TOTAL			**35%**	**1017**	**54.71**
Buchanan	2003-04	11	10	91%	30	3
	2004-05	11	0	0%	0	0
	2005-06	10	0	0%	0	0
	2006-07	12	4	33%	21	5.25
	2007-08	12	0	0%	0	0
	TOTAL			**25%**	**51**	**8.25**
Cambridge	2003-04	14	9	64%	39	4.33
	2004-05	15	7	47%	176	25.14
	2005-06	14	13	93%	173.95	13.38
	2006-07	15	15	100%	379.5	25.3
	2007-08	15	2	13%	18.5	9.25
	TOTAL			**63%**	**786.95**	**77.41**
Carter	2003-04	5	1	20%	36	36
	2004-05	6	1	17%	72	72
	2005-06	6	6	100%	90	15
	2006-07	6	0	0%	0	0
	2007-08	6	2	33%	24	12
	TOTAL			**34%**	**222**	**135**
Carver	2003-04	17	13	76%	139.25	10.71
	2004-05	17	15	88%	566.75	37.78
	2005-06	18	18	100%	591.25	32.85
	2006-07	18	2	11%	76	38
	2007-08	16	0	0%	0	0
	TOTAL			**55%**	**1373.25**	**119.34**

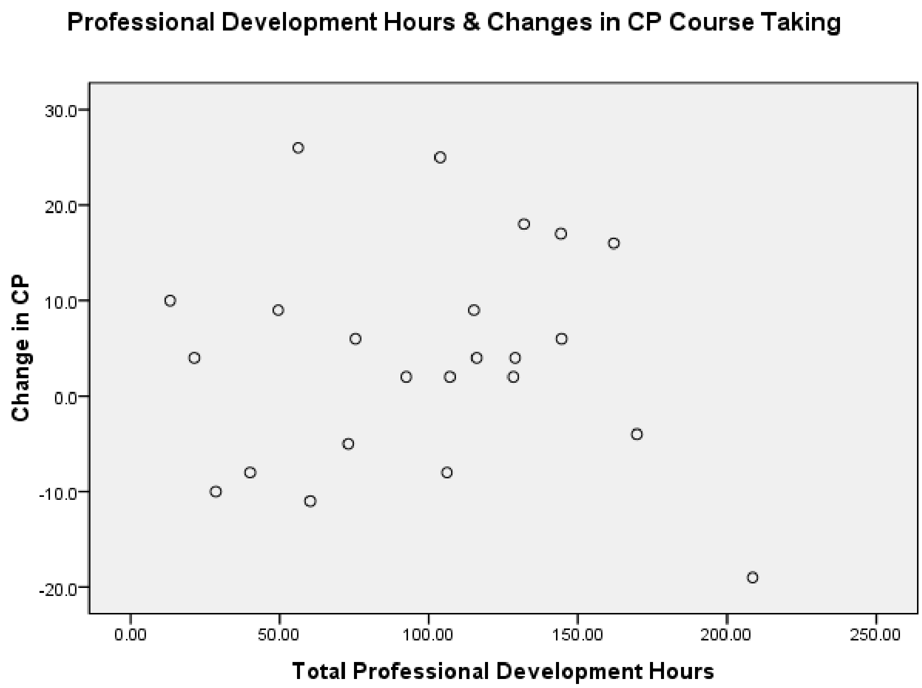
School District	*Academic Year*	*Total Number of Eligible Teachers*	*Total Number of Participating Teachers*	*Percent of Eligible Teachers who Participated*	*Total Professional Development Hours*	*Average Professional Development Hours*
Erikkson	2003-04	13	2	15%	26	13
	2004-05	18	15	83%	150	10
	2005-06	18	18	100%	313	17.39
	2006-07	16	15	94%	73.25	4.88
	2007-08	14	5	36%	50.75	10.15
	TOTAL			**66%**	**613**	**55.42**
Fairlea	2003-04	23	0	0%	0	0
	2004-05	23	3	13%	90	30
	2005-06	21	4	19%	80	20
	2006-07	24	0	0%	0	0
	2007-08	23	4	17%	45.5	11.38
	TOTAL			**10%**	**215.5**	**61.38**
Grant	2003-04	12	2	17%	32	16
	2004-05	13	6	46%	199.5	33.25
	2005-06	14	10	71%	270	27
	2006-07	13	9	69%	279	31
	2007-08	13	1	8%	2.5	2.5
	TOTAL			**42%**	**783**	**109.75**
Hayes	2003-04	30	0	0%	0	0
	2004-05	29	16	55%	43.5	2.72
	2005-06	27	7	26%	138	19.71
	2006-07	30	2	7%	59	29.5
	2007-08	30	9	30%	120	13.33
	TOTAL			**24%**	**360.5**	**65.27**
Jackson	2003-04	16	0	0%	0	0
	2004-05	17	12	71%	75	6.25
	2005-06	17	1	6%	12	12
	2006-07	18	0	0%	0	0
	2007-08	17	0	0%	0	0
	TOTAL			**15%**	**87**	**18.25**

School District	*Academic Year*	*Total Number of Eligible Teachers*	*Total Number of Participating Teachers*	*Percent of Eligible Teachers who Participated*	*Total Professional Development Hours*	*Average Professional Development Hours*
Jericho Township	2003-04	7	2	29%	80	40
	2004-05	9	6	67%	229	38.17
	2005-06	10	0	0%	0	0
	2006-07	11	6	55%	182	30.33
	2007-08	8	0	0%	0	0
	TOTAL			**30%**	**491**	**108.5**
Madison	2003-04	21	1	5%	3	3
	2004-05	17	1	6%	49.5	49.5
	2005-06	16	4	25%	39.75	9.94
	2006-07	16	1	6%	6.5	6.5
	2007-08	16	0	0%	0	0
	TOTAL			**8%**	**98.75**	**68.94**
Midvale	2003-04	8	2	25%	66	33
	2004-05	9	6	67%	36	6
	2005-06	9	1	11%	4	4
	2006-07	8	0	0%	0	0
	2007-08	9	0	0%	0	0
	TOTAL			**21%**	**106**	**43**
North Ford	2003-04	17	0	0%	0	0
	2004-05	17	16	94%	89.25	5.58
	2005-06	17	8	47%	198	24.75
	2006-07	18	18	100%	175.5	9.75
	2007-08	16	0	0%	0	0
	TOTAL			**48%**	**462.75**	**40.08**
Pierce	2003-04	22	6	27%	113	18.83
	2004-05	21	17	81%	262	15.41
	2005-06	23	23	100%	461	20.04
	2006-07	22	22	100%	278	12.64
	2007-08	23	0	0%	0	0
	TOTAL			**62%**	**1114**	**66.92**

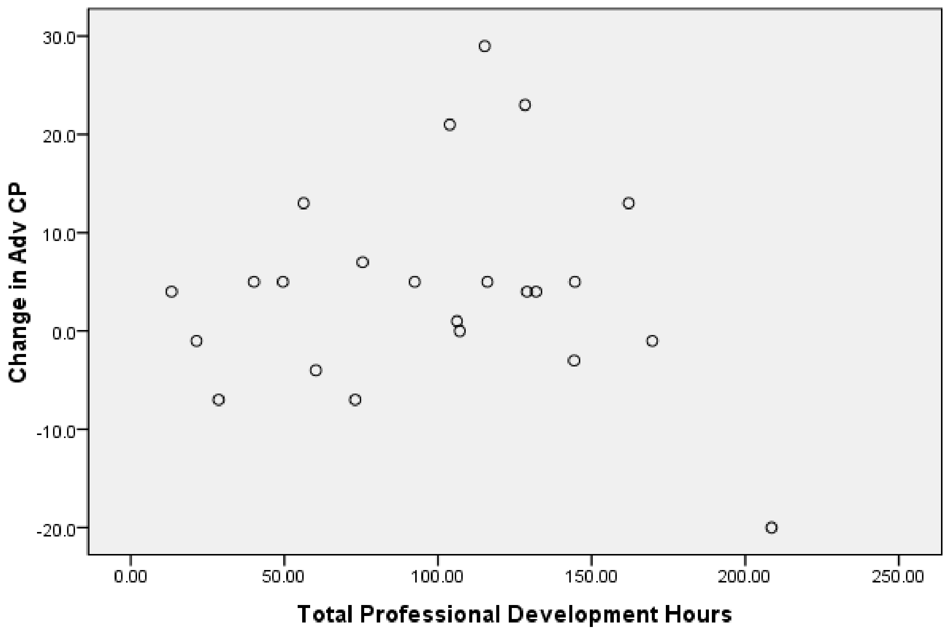
School District	*Academic Year*	*Total Number of Eligible Teachers*	*Total Number of Participating Teachers*	*Percent of Eligible Teachers who Participated*	*Total Professional Development Hours*	*Average Professional Development Hours*
Rittenhouse	2003-04	8	4	50%	10	2.5
	2004-05	8	6	75%	185.25	30.88
	2005-06	8	8	100%	188	23.5
	2006-07	8	8	100%	48	6
	2007-08	8	0	0%	0	0
	TOTAL			**65%**	**431.25**	**62.88**
Roosevelt	2003-04	17	7	41%	196	28
	2004-05	18	7	39%	90	12.86
	2005-06	19	4	21%	103	25.75
	2006-07	18	0	0%	0	0
	2007-08	18	0	0%	0	0
	TOTAL			**20%**	**389**	**66.61**
Rye	2003-04	21	0	0%	0	0
	2004-05	24	24	100%	229.5	9.56
	2005-06	25	20	80%	90	4.5
	2006-07	25	2	8%	6	3
	2007-08	26	2	8%	7.5	3.75
	TOTAL			**39%**	**333**	**20.81**
South Harding	2003-04	14	0	0%	0	0
	2004-05	14	8	57%	108	13.5
	2005-06	13	1	8%	18	18
	2006-07	14	0	0%	0	0
	2007-08	15	0	0%	0	0
	TOTAL			**13%**	**126**	**31.5**
Taft	2003-04	13	0	0%	0	0
	2004-05	13	0	0%	0	0
	2005-06	14	14	100%	138.25	9.88
	2006-07	13	0	0%	0	0
	2007-08	13	0	0%	0	0
	TOTAL			**20%**	**138.25**	**9.88**

School District	*Academic Year*	*Total Number of Eligible Teachers*	*Total Number of Participating Teachers*	*Percent of Eligible Teachers who Participated*	*Total Professional Development Hours*	*Average Professional Development Hours*
Taylor	2003-04	17	0	0%	0	0
	2004-05	17	17	100%	537.5	31.62
	2005-06	16	16	100%	385.5	24.09
	2006-07	18	18	100%	357.5	19.86
	2007-08	17	9	53%	49.5	5.5
	TOTAL			**71%**	**1330**	**81.07**
Twain	2003-04	8	8	100%	101	12.63
	2004-05	10	5	50%	24.5	4.9
	2005-06	11	11	100%	103.5	9.41
	2006-07	10	0	0%	0	0
	2007-08	11	0	0%	0	0
	TOTAL			**50%**	**229**	**26.93**
Tyler	2003-04	23	7	30%	112	16
	2004-05	24	17	71%	394	23.18
	2005-06	22	2	9%	21.25	10.63
	2006-07	24	2	8%	32	16
	2007-08	25	1	4%	18	18
	TOTAL			**25%**	**577.25**	**83.8**

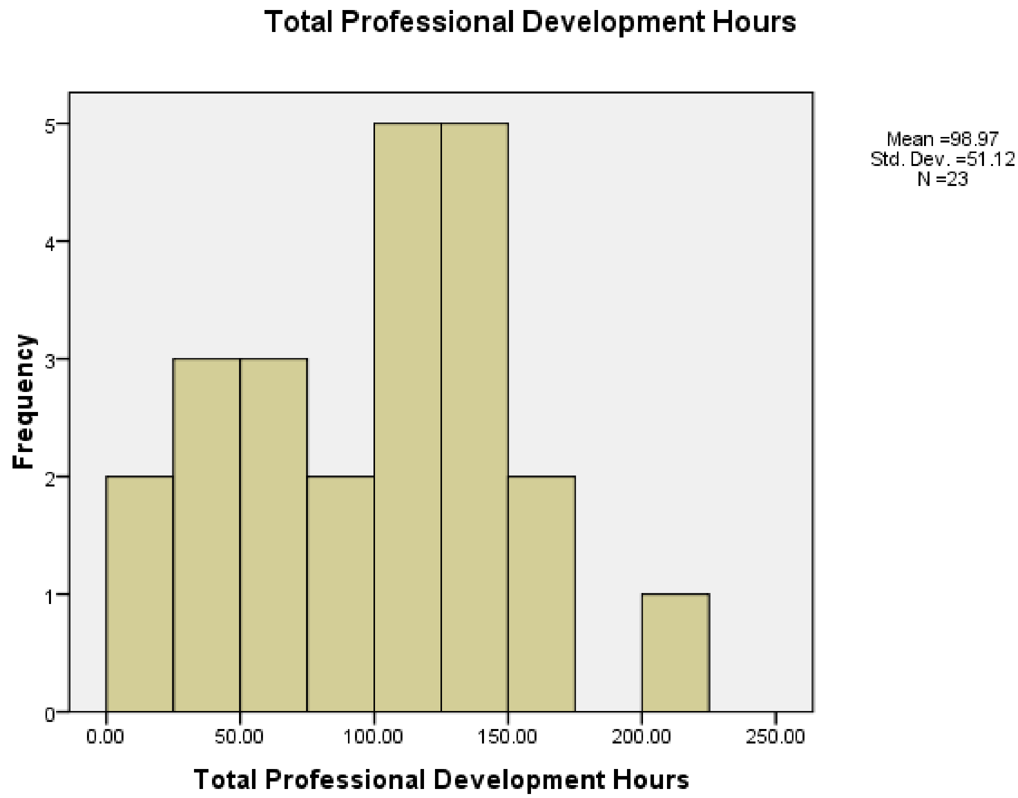
APPENDIX E
SCATTER PLOTS: PROFESSIONAL DEVELOPMENT AND CHANGES IN MATHEMATICS OUTCOMES

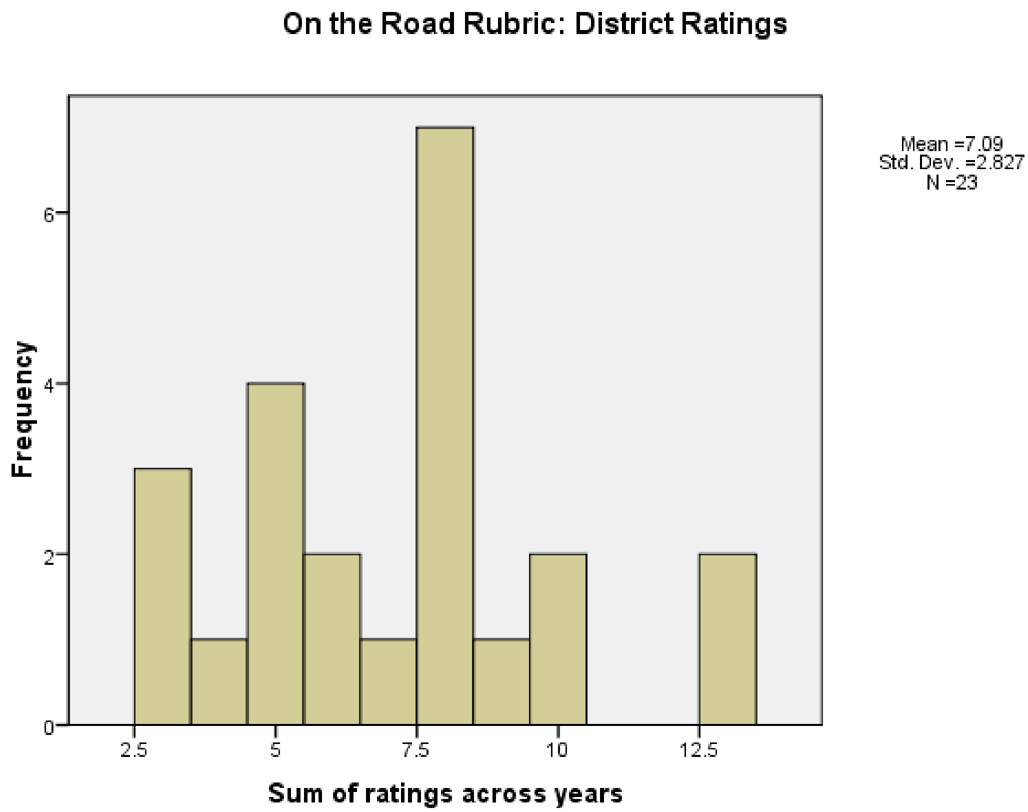

Professional Development Hours & Changes in CP Course Taking
Change in CP
30.0
20.0
10.0
0.0
-10.0
-20.0
0.00
50.00
100.00
150.00
200.00
250.00
Total Professional Development Hours

Professional Development Hours & Changes in Advanced CP Course Taking



APPENDIX F
HISTOGRAMS OF DICHOTOMOUS VARIABLES




On the Road Rubric: District Ratings
Mean =7.09
Std. Dev. =2.827
N =23
Frequency
0
2
4
6
2.5
5
7.5
10
12.5
Sum of ratings across years