WORKING
PAPER

Math Science Partnership of Southwest Pennsylvania

Year Four Evaluation Report

CYNTHIA A. TANANIS
JOHN F. PANE
VALERIE L. WILLIAMS
STUART S. OLMSTED
CARA CIMINILLO
SHANNAH THARP-TAYLOR

WR-494-AIU

June 2007

Prepared for Allegheny Intermediate Unit

EDUCATION

Preface

In 2002, the National Science Foundation (NSF) launched the Math and Science Partnership program. This program builds on the nation's dedication to educational reform by supporting partnerships that unite the efforts of K-12 school districts with science, mathematics, engineering, and education faculties of colleges and universities. This program has made 80 awards with over $500 million awarded to date, not including budgeted funds for future years in multi-year projects. The Math Science Partnership of Southwest Pennsylvania (MSP) is one of seven comprehensive partnership projects funded by NSF in 2003. This MSP brings together 53 K-12 school districts – 45 as part of the NSF grant as well as eight additional districts supported by a companion Math Science Partnership grant from the Pennsylvania Department of Education – four Intermediate Units (IUs), four Institutions of Higher Education (IHEs), and other strategic partners in Southwest Pennsylvania. The goals are to increase K-12 students' knowledge of mathematics and science; increase the quality of the K-16 educator workforce; and create a sustainable coordination of partnerships in the IUs. The MSP is housed at the Allegheny Intermediate Unit (AIU), in Homestead, Pennsylvania near Pittsburgh. AIU subcontracted with the RAND Corporation and the University of Pittsburgh to serve as the project's evaluation team. The project and the evaluation commenced in September 2003.

The evaluation investigates the effectiveness of the partnership, its impact on institutional practices and policies at partner educational institutions, changes in math and science instruction, and changes in student course taking and outcomes. Over the course of the project, data will be collected from numerous sources to address these points, including focus groups and interviews of key project personnel, surveys of principals and math and science teachers, case studies in partnership school districts and IHEs, documentation of partnership meetings and activities, artifacts produced by the partnership, math and science achievement data for K-12 students, and course completion data for K-12 and IHE students.

This working paper is based on information collected from the project's start in September 2003 through December 2006. It is the fourth in a series of annual evaluation reports that the Assessment and Evaluation Team has provided to AIU, which partially fulfill AIU's larger annual reporting requirements to the NSF. In contrast to previous annual evaluation reports that described evaluation progress-to-date, the primary purpose of the Year 4 report is to examine the MSP in terms of what seems to be working and to develop hypotheses about the circumstances under which this MSP may be successful.

The study was conducted by RAND Education, a unit of the RAND Corporation, and the University of Pittsburgh's School of Education, under contracts with the Allegheny Intermediate Unit.

Table of Contents

Figures

Tables

Summary

The Math Science Partnership of Southwest Pennsylvania (MSP) is a collaboration among 53 K-12 school districts, four Intermediate Units (IUs),[1] four Institutions of Higher Education (IHEs), and other strategic partners in Southwest Pennsylvania. The goals of the partnership are to increase K-12 students' knowledge of mathematics and science; increase the quality of the K-16 educator workforce; and create a sustainable coordination of partnerships in the IUs. RAND, in collaboration with the University of Pittsburgh, is conducting a five-year evaluation of the partnership. The evaluation is measuring the effectiveness of the partnership, its impact on institutional practices and policies at partner educational institutions, changes in math and science instruction, and changes in student outcomes. This document represents the Year Four evaluation report. It is intended to lay groundwork for the evaluation team's culminating final assessment.

Focus on Participation

This report differs from the earlier annual evaluation reports, which focused primarily on analyzing data about intervention activities, reform implementation, and interim progress toward goals. For this year's report, the evaluation team chose instead to use participation levels as an analytical focus to identify promising practices among MSP partners and to use only a subset of MSP data to zero in on the impact of participation.

The rationale for this choice is as follows: *Participation* in MSP activities should lead to *implementation* of reformed policies and practices, which in turn lead to *attainment* of project's goals. Although most of the evaluation data collection focuses ultimately on measuring outcomes, participation is a crucial first step toward desired outcomes.

The MSP intervention includes professional development intended to enhance content knowledge and leadership, support the alignment of curriculum and pedagogy with reform goals, and disseminate and support the use of research-based resources and tools. The impact of the MSP on both mid- and long-term outcomes will likely be stronger in districts that participate more in these activities and weaker in districts that participate less. Measuring participation is therefore central to this evaluation for two reasons. First, participation may help to explain differences in observed effects across districts. Second, correlation between participation and outcomes can support a case that the MSP is responsible for those changes.

The team defined measures of participation for four of the evaluation's main data collections: (1) For K-12 district case studies, the evaluation team chose a measure of staff participation: *total hours adjusted for staff size*, or THASS--to contrast high- versus low-participating districts. (2) For principals surveyed in each of the 48 districts, the team supplemented THASS with another proxy for MSP participation, principals' involvement in *Lenses on Learning*, a mathematics professional development course for school leaders. (3)

[1] IUs are publicly funded educational service agencies, serving as regional intermediaries between local control school districts and the Pennsylvania Department of Education.

For student achievement, the team used the THASS variable and measured its relationship to various indicators of student performance. (4) For IHEs, the evaluation team identified faculty members as high-participating, based on the estimated total hours of participation. For each of these areas, the team identified changes observed since the beginning of the MSP and explored their relationship to participation levels. The key findings appear below.

Key Findings

K-12 school district case studies

In high-participation districts, there were observed changes in curriculum and teaching practice. Observations and interview data suggested increased teacher understanding and awareness of the value of research-based teaching practices and materials, which in turn have affected curriculum choices. In addition, many professional development participants felt they possessed deeper content knowledge in math and science.

Observations indicated that math teachers are benefiting from MSP activities that focus on the use of active learning, and are pleased with what they are learning. Moreover, in all but one of the case study classroom observations in high-participating districts, elementary math instruction was rated as *effective* to *exemplary* and included active student learning opportunities. Generally, classroom observations in districts with less participation yielded somewhat lower ratings.

Case-study observations tended to rate science classroom instruction somewhat lower than math. While instruction was typically purposeful and well-planned, implementation of the lessons seemed more challenging. Lessons were often "short circuited" by a lack of time. However interviews suggested that teachers have appreciated and benefited from the Teacher Leadership Academies (TLAs) and on-site academies.

Some of these changes may be attributable to the MSP. Teacher leaders and some teachers who have received training at on-site academies have demonstrated changes in the way they approach what they teach as well as the way they implement lessons.

Survey of Principals

The evaluation team surveyed principals in 2004 to collect baseline data on principals' views, attitudes, policies, and practices on a range of subjects related to the MSP. In 2006, the evaluation team administered the same survey to assess principals' responses after further implementation of the MSP project.

The evaluation team observed several changes from the baseline survey. In the 2006 survey, principals reported higher levels of agreement that they are knowledgeable about national standards for science, knowledgeable about the district science curriculum, and prepared to support teachers in implementing the science curriculum. They also reported an increase on a parallel set of mathematics questions,

although this difference was not significant. The principals reported being more comfortable serving as math and science instructional leaders, including discussing concrete examples of instructional practices with teachers, observing classrooms, examining student work, and providing feedback on teaching. Finally, principals responded more strongly that developing a professional learning community—giving teachers time to prepare lessons, work with other teachers, attend professional development—encourages effective instruction.

Principals' responses concerning district and school policies and practices also changed significantly. On two scales, which include items such as the regional curriculum framework, testing policies, access to computers and calculators, and the quality of district adopted instructional materials, principals reported higher on the scale that these items "encourage effective instruction". The effect size for these changes is greater than those associated with leadership skills.

Some of these changes may be attributable to participation in the MSP—specifically, to *Lenses on Learning*. Principals who had taken the *LoL* training placed a higher value on mathematics and science instruction, the effect of quality resources, and the importance of supporting continuous professional development for their teachers. These principals also reported greater commitment to providing a supportive environment for reform-based pedagogy. Finally, principals who have attended *LoL* said they held a deeper understanding of what constitutes good mathematics and science lessons and a greater confidence and ability to support them.

Student outcomes

Student achievement is the ultimate measure of the MSP's impact. Outcomes data through Year Four of the evaluation have not shown significant relationships between MSP participation and student math achievement. However, benchmark and other data show encouraging changes in short-term and mid-term outcomes. Such changes are generally accepted as necessary precursors to changes in achievement, although for the most part these changes cannot be causally attributed to the MSP. Additional data from Year Five will enable recalculating achievement models for more cohorts of students, in science as well as math, over more years of project implementation, and with a fourth year of participation data.

IHE partners

A number of findings related to the MSP were observed among IHE faculty. These centered on building partnerships with K-12 educators, classroom teaching, and planning classes.

Many high-participation faculty members, especially those involved in TLAs and the math/science leadership teams, reported strong relationships with the MSP Coordinators and other MSP staff, who, as employees of the IUs, are K-12-based.

High-participating faculty members also reported that attending TLAs helped them to become better teachers. These faculty members benefit from participating in the MSP through improvements in the quality of their teaching. This benefit derives, at least in part, from teaching being valued at their institution and counted towards professional advancement. However, respondents also noted that some

teachers who implemented teaching practices promoted by the MSP saw their student evaluation scores decline, possibly due to the additional demands placed on students. Therefore, although MSP participation may lead to improved pedagogy, the potential for poor student evaluations should be considered by the administration when making tenure considerations.

High-participating faculty members also noted the compatibility of MSP goals and their IHE's goal to provide a strong teaching certification program. Many respondents recognized that participation in the MSP had afforded their institution's teacher certification programs more regional recognition among K-12 districts, and more national recognition among peer institutions. This recognition of the value of MSP involvement facilitated faculty participation in MSP activities.

Emerging Themes

Several cross-cutting themes related to high participation in the MSP emerged from the Year Four analysis.

Curriculum and teaching. A common thread across the data from K-12 districts and the IHEs is that changes in the curriculum and in teaching practice have occurred in high-participation partners and that some may have resulted from the MSP. Teacher leaders and some teachers who have received training at on-site academies have demonstrated changes in the way they approach what they teach as well as the way they implement lessons. Further, principals report a deeper understanding of what constitutes good mathematics and science lessons; and IHE faculty report changes in how they plan and deliver their courses.

Time management. The MSP is a long-term, time-intensive intervention that involves considerable training and professional development time. Time management issues thus pose a substantial concern. High-participation districts appear to have been more successful in meeting the time management challenge posed by the MSP's professional development requirements. Schools nonetheless struggle to enact the on-site professional development, in part due to reductions in the number of hours available for training, and planned sessions are often supplanted by competing priorities and professional development needs. IHE partners struggle with time management in a somewhat different way. Scholarship and teaching are more highly valued in the promotion and tenure process and thus they struggle to find time for MSP activities.

Leadership and advocacy. There has been consistent evidence across the first four years of MSP implementation that some advocacy is needed in schools to champion a reform initiative such as the MSP. At times this advocacy has started with teachers, at other times, with administrators. However, there is also evidence to suggest that administrative support, and perhaps, advocacy, is required for such an intensive effort to be sustainable. At the IHE, similar issues emerge. The evaluation team has reported in the past that partnership, especially at the IHE level, is forged first and foremost, via individual faculty. While an institution may indicate a policy of partnership in the MSP, this is often translated through a few faculty members who advocate for the efforts. Across the four years of the MSP to date, there has been considerable effort by the Deans of participating IHEs to forge relationships with one another and

articulate their support of the MSP and its sustainability. One example of this has been the Dean's Dinners and a joint policy statement regarding the importance of faculty involvement, especially in relation to promotion and tenure. Even with these efforts, planning for sustainability among the IHE partners seems less tangible among faculty than among some K-12 school and IU partners.

Adaptation and sustainability. The annual MSP implementation plans have acknowledged the need to promote sustainability among partners beyond the five-year NSF funding period. Over time, K-12 and IHE partner institutions change, as do priorities and resources. K-12 changes sometimes occur rapidly. This is exacerbated by staff turnover or changes in teaching assignments, diluting the training effect within a particular grade level, discipline, or school. Districts' adaptation of the implementation plan helps to address these issues. However, the project and the schools struggle to balance "acceptable" versus "fatal" adaptation for meeting the goals of the project at particular sites.

Higher education faculty work more independently than do K-12 teachers. Consequently, adaptation is less problematic for individual faculty at the IHE level. Adaptations to curricula or pedagogy are much more likely to stay with that particular individual and there are fewer opportunities for dissemination across colleagues. Moreover, higher education faculty do not as often work from a common syllabus and incoming faculty members may choose to teach courses quite differently than their predecessors.

Cultural issues, both in K-12 and IHE, also present challenges in adaptation and sustainability across the partnership. The K-12 institutional environment is designed to be amenable to public policy change through state and local control. There has been visible evidence of how much both federal legislation and state mandates impact how education is organized and delivered as a result of various state standards and assessments and other policies stemming from No Child Left Behind. The more recent flurry of educational accountability measures has created an environment where change is expected. This is less evident in IHEs currently, though new efforts are afoot to bring similar measures of accountability to bear in post-secondary education.

The achievement horizon. Evidence is strong that change is occurring among K-12 educators and IHE faculty. However, it remains unclear how long it may take for these changes to affect student achievement. Many questions regarding the achievement horizon remain, notably whether factors beyond the reach of a reform such as the MSP are pushing the horizon further away despite best efforts. The evaluation's final report will attempt to shed more light on this issue.

Acknowledgements

The authors would like to thank David Adamson, Crystal Baksis, Valerie Becker, Christopher Beighley, Sarah Hooper, Dan McCaffrey, Deb Moore, Nancy Olenik, Caleb Smith, and Keith Trahan for their contributions to this report. We would also like to acknowledge the MSP staff, IHE faculty and administrators, and the K-12 teachers that participated in our data collection activities. Finally, we would like to thank Nancy Bunt, the principal investigator of the Math Science Partnership of Southwest Pennsylvania for the helpful comments we received on this report. This research was performed under an agreement funded by the National Science Foundation; award number EHR-0314914. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the project leadership, the Allegheny Intermediate Unit, the National Science Foundation or the U.S. Government.

Acronyms

Acronym	Definition
AIU	Allegheny Intermediate Unit
CDS	Content Deepening Seminar
IHE	Institution of Higher Education
IU	Intermediate Unit
LAA	Leadership Action Academy
LAT	Leadership Action Team
LoL	Lenses on Learning
MSP	Math Science Partnership of Southwest Pennsylvania
NCLB	No Child Left Behind
NSF	National Science Foundation
PDE	Pennsylvania Department of Education
PI	Principal Investigator
PSSA	Pennsylvania System of School Assessment
RETA	Research, Evaluation, and Technical Assistance
STEM	Science, Technology, Engineering, and Mathematics
TIMSS	Trends in International Mathematics and Science Study
TLA	Teacher Leadership Academy

1. Introduction

This report is the fourth in a series of annual evaluation reports that detail the evolution of the Math Science Partnership of Southwest Pennsylvania (MSP). In contrast to previous annual evaluation reports that described evaluation progress-to-date, the primary purpose of the Year 4 report is to examine the MSP in terms of what seems to be working and to develop hypotheses about the circumstances under which this MSP may be successful. Whereas the previous annual evaluation reports were more comprehensive in the amount of data discussed, the Year 4 report focuses on a subset of data and explores the outcomes based on the participation levels of MSP partners. This approach is taken in preparation for the Year 5 Summative Evaluation report – by developing hypotheses at this stage, the evaluation team anticipates being able to address questions of whether the MSP has achieved its goals and also questions about the contextual factors that either facilitated or hindered the achievement of the project goals.

Overview of the Math Science Partnership Program

The National Science Foundation's (NSF) Math Science Partnership (MSP) program was an outgrowth of the No Child Left Behind (NCLB) Act of 2001, focusing on three issues: an excess of teachers teaching out of field; too few students taking advanced coursework; and too few schools offering challenging curricula and textbooks. Although previous NSF programs targeting math and science educational reform in levels K – 12 have focused on similar issues, the Math Science Partnership program also includes higher education as a critical partner and potential benefactor of its efforts.

The Math Science Partnership program supports two types of partnerships: *comprehensive* and *targeted*. The comprehensive projects are funded for a five-year period for up to $7 million annually. These projects are intended to implement change in mathematics and/or science education in both institutions of higher education (IHEs) and school districts, resulting in improved student achievement across the K-12 and IHE continuum. The targeted projects focus on improved student achievement in a narrower grade range within K-12 or with a focus on a single discipline (math or science), and are funded for up to $2.5 million per year for up to five years. In addition, the Math Science Partnership program funds research, evaluation, and technical assistance (RETA) projects that build and enhance large-scale research and evaluation capacity for all Math Science Partnership awardees and provide them with tools and assistance in the implementation and evaluation of their work.

NCLB also authorizes a parallel Math Science Partnership program at the U.S. Department of Education. This program requires partnerships to include a state educational agency or a public regional intermediary, such as one of Pennsylvania's Intermediate Units (IUs), to work with the engineering, math or science department of an institution of higher education, and a high-needs school district. Unlike the NSF program, where funds are awarded in a national competition, the Math Science Partnership program at the Department of Education awards funds to states to administer.

The Math Science Partnership of Southwest Pennsylvania

The Math Science Partnership of Southwest Pennsylvania (SW PA MSP) is one of seven comprehensive partnership projects funded by NSF in 2003. It is a partnership of 53 school districts, four institutions of higher education (IHEs), and four partner Intermediate Units (IUs).[2] The NSF award has supported 40 of the school districts since the project's inception, and another five were added in year 4 (2006-07) as part of a planned expansion. A Math Science Partnership award in the U.S. Department of Education program, through the Pennsylvania Department of Education (PDE), supports the remaining eight districts. The MSP is headquartered at the Allegheny Intermediate Unit (AIU) near Pittsburgh; this IU serves the greatest density of participating/partner school districts in the region.

The region of Southwestern Pennsylvania that is served by the MSP includes the urban fringe of the city of Pittsburgh, several smaller urban areas, suburbs, and rural areas.[3] Total enrollment in the MSP school districts is approximately 114,000 students, with approximately 3,800 teachers who teach math or science topics. On average, about 39% of students in MSP schools are economically disadvantaged,[4] compared with a statewide average of 36%. This figure is higher in the PDE MSP districts (59%) than in the NSF MSP districts (35%). The enrollment of underrepresented minorities[5] is approximately 19%, compared with a statewide average of 22%. Again, this figure is higher in the PDE MSP districts (25%) than in the NSF MSP districts (18%). These demographics vary widely across schools. The reported percentages for both economically disadvantaged and minority populations vary from 0% to nearly 100% in individual schools.

Similarly, there is wide variation in student achievement levels across the K-12 schools in the MSP. A substantial portion of MSP schools are not making adequate yearly progress under NCLB; three MSP districts are identified as "empowerment districts" meaning that they are subject to state control if they do not improve, and one of those districts is already being operated under a state board of control. At the other end of the spectrum, the MSP includes several "blue ribbon" schools, which are among the highest achieving in the state.

The four partner IHEs are small- to mid-sized, teaching-oriented, private institutions located in southwest Pennsylvania: Carlow University, Chatham University, Robert Morris University, and Saint Vincent College. Approximately 8,600 students are enrolled in these IHEs, and approximately 59 members of their math, science, engineering and education faculties are participating in this project. Although some of the

---

[2] Intermediate Units are publicly funded educational service agencies that act as regional intermediaries between local school districts and the Pennsylvania Department of Education.

[3] Pittsburgh Public Schools (PPS), the largest urban school district in the region, is not formally involved as an MSP participant.

[4] As is common practice, we use free or reduced-price lunch eligibility as a proxy for economic status.

[5] The racial/ethnic groups included in this category are African-American, Hispanic, Asian, and Native American students.

larger, research-oriented universities in southwest Pennsylvania were invited to participate in the MSP, they declined. In some cases, the universities were already involved in educational reform programs. For example, the University of Pittsburgh School of Education was already involved in a Math Science Partnership through the University's Learning Research and Development Center.

Consistent with the objectives of the overall Math Science Partnership program, the fundamental goals of the SW PA MSP are to (1) increase K-12 students' knowledge of mathematics and science, (2) increase the quality of the K-16 educator workforce, and (3) create a sustainable coordination of partnerships in the IUs, building intentional feedback loops between K-12 districts and IHEs, tapping the discipline-based expertise of the IHEs, and improving the mathematics and science learning experiences for all undergraduates.

Evaluation Design and Implementation

The AIU subcontracted with the University of Pittsburgh, the RAND Corporation, and the Evaluation, Grants and Data Division of AIU to serve as evaluators for the SW PA MSP. Collectively, individuals from these organizations form the evaluation team, which monitors the project's progress in order to offer formative advice to the project, measures its ultimate success in achieving its goals, and documents lessons for the benefit of future initiatives that may seek to replicate it. Four research questions guide the evaluation:

1. Have MSP partners developed and implemented a comprehensive intervention targeting math and science curriculum and achievement? If so, how?

2. Have institutional practices and support structures changed at K-12 districts and IHEs participating in the MSP? If so, how?

3. Has math and science instruction changed in K-12 districts participating in the MSP? If so, how?

4. In what ways have student outcomes and course taking changed in K-12 schools and districts participating in the MSP? If change occurred, what is the connection between implementation of the MSP plan and these changes?

To address these research questions, the evaluation team adopted a mix of qualitative and quantitative methods in three distinct but overlapping areas of research and analysis: (1) formative assessment and documentation of MSP activities in relation to the institutional goals and student outcomes described above; (2) qualitative and quantitative investigation of implementation at K-12 districts, including (a) institutional change at the district level, and (b) the links between involvement in partnership activities and curriculum implementation strategies at the district and school level and K-12 student outcomes; and (3) evaluation of institutional change at IHE partners as a result of involvement in MSP activities.

Table 1-1 shows the evaluation strategies, data sources, and sampling methods used in the evaluation, and indicates which research questions are informed by those sources. The table also indicates how these data collection strategies are related to the research questions and how frequently the data are being

collected. Full copies of the protocols and instruments are not included in this report; however, they are available to project staff and other interested parties on request.

Table 1-1: Evaluation Study Design

Evaluation Strategy	Data Source	Research Questions Addressed	Frequency of Data Collection	Type of Information Gathered	Sampling Method
Observations	MSP Events	1, 2	Years 1-4	Implementation data	Representative set of major project activities
Interviews / Focus Groups	Project PI and MSP Coordinators	1, 2	Years 1-4	Implementation data	All
Interviews / Focus Groups	IHE Faculty	1, 2	Years 1-4	Implementation data	Representative sample
Case studies	School districts	1, 2, 3	Years 2-4	Implementation and Impact data	Purposive sample of school districts
Case Studies	IHEs	1, 2	Years 4 & 5	Implementation and Impact Data	All IHEs
Survey	Teachers	2, 3	Years 1 & 4	Impact data	Random sample of teachers
Survey	Principals	2, 3	Years 2 & 4	Impact data	All principals
LMT (Math Content Knowledge)	Teachers	3	Years 3-5	Impact data	Teacher leaders and some teachers (selected via academy enrollment)
Pre-post and statewide comparisons	Student achievement data	4	Years 1-4	Impact data	K-12 students in tested grades
Pre-post and regional comparisons	Course completion data	4	Years 1-4	Impact data	K-12 graduates

SW PA MSP Logic Model and Theory of Action

To assist in the evaluation, the evaluation team developed a logic model to illustrate the interrelationships among the SW PA MSP program's goals, activities, and outcomes. Logic models are common evaluation tools, and offer visual representations of a program's path to achieving its intended outcomes. The model not only describes the MSP project, but also furnishes the evaluation team with a unified set of terms and relationships to facilitate discussion and thinking about the MSP project. The logic model is viewed as a work in progress that evolves with the evaluation team's thinking and analysis of the MSP project. The current version of the logic model is shown in Figure 1-1. It includes the traditional components of inputs, activities, outputs, and outcomes.

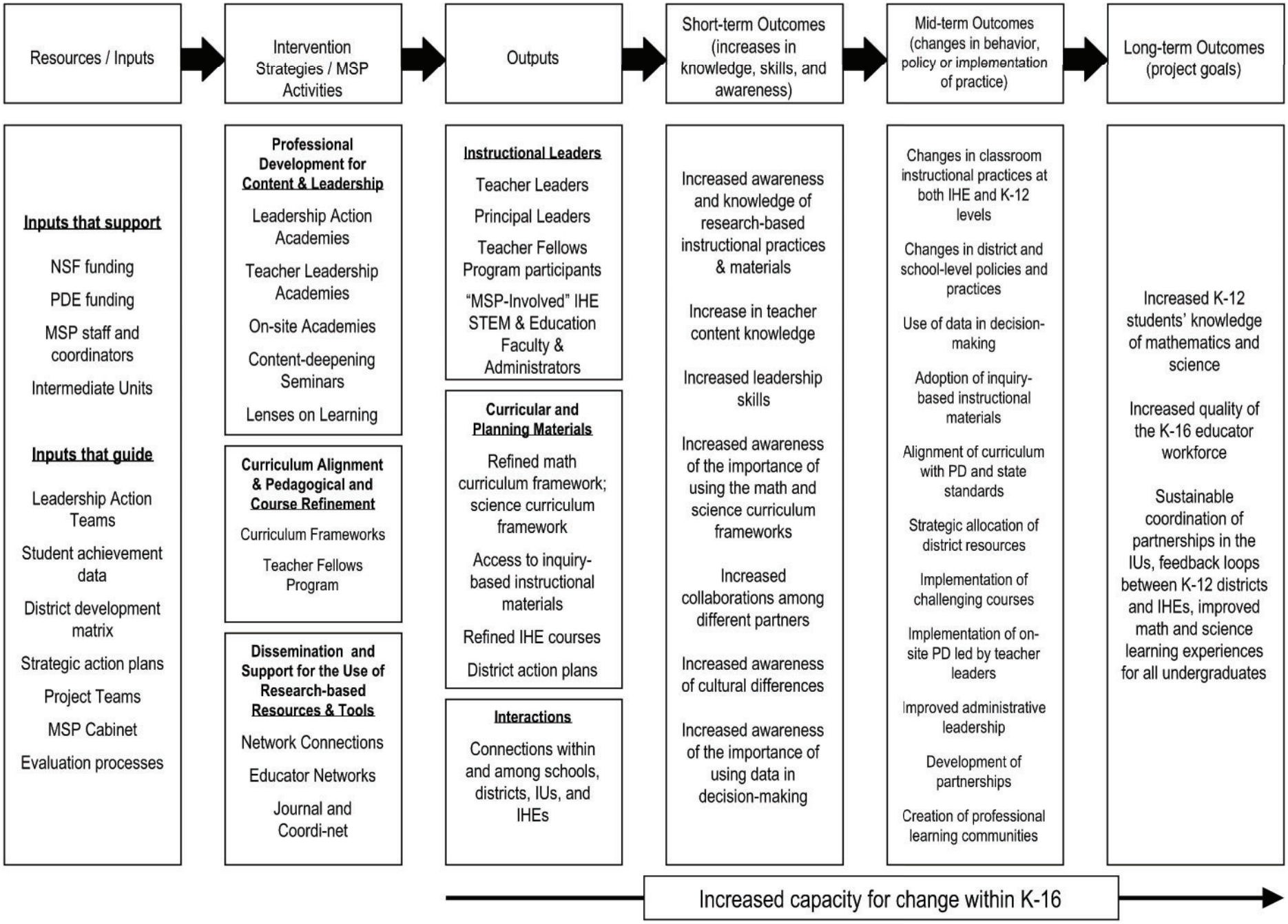


Figure 1-1: SW PA MSP Logic Model

Theory of Action

The theory of action that underlies the MSP logic model is premised on the view that student achievement in mathematics and science can be enhanced by administrators and classroom teachers who are willing to become learners and deepen their own conceptual understanding of the big ideas in mathematics and science.[6] Similar to the theory of action for the NSF-supported Local Systemic Change program, this theory of action argues that providing teachers with opportunities to deepen their content and pedagogical knowledge in the context of high-quality instructional materials will result in better prepared teachers. With ongoing support, teachers will be more inclined to change their instruction in ways advocated by national standards, and will have more capacity to do so. These changes in instructional practices are expected to improve student thinking and learning, a core focus of the MSP, which will in turn lead to higher student achievement.

This theory of action is supported by a number of research studies. In particular, Catherine Lewis's seminal work on "lesson study" and research by Japanese and U.S. educators has outlined key pathways to instructional improvement, most of which are mirrored in the MSP logic model. These key pathways include: increased knowledge of subject matter; increased knowledge of instruction; increased ability to observe students; stronger collegial networks; stronger connection of daily practice to long-term goals; strong motivation and sense of efficacy; and improved quality of lesson plans (Lewis et al., 2004). The intervention strategies that the MSP has employed provide mechanisms for achieving program goals along each of these pathways. The teacher leadership academies develop understanding of content and pedagogy and are important pathways for achieving increased knowledge of instruction as well as content. Content deepening seminars are designed to update and deepen teacher knowledge in a specific content area and contribute to increased knowledge of subject matter. Stronger collegial networks are built via a number of routes, including participation in Network Connections[7]; district led on-site PD academies; and Educator Networks. The emphasis on curriculum alignment and pedagogical and course refinement offers opportunities to improve the quality of instructional materials and lesson plans. However, in addition to these pathways, the MSP theory of action also argues that support from district leadership is an important component of instructional improvement. Thus, *Lenses on Learning* is a key intervention strategy for gaining administrator support. In order for administrators to support effective teaching and learning in the classroom, they must first learn what good instruction is and how to recognize it. With these new "lenses," administrators are better prepared to support teacher-led instructional change, through improved teacher observation skills, support for professional development, and the creation of strong learning communities within the schools and districts. Finally, the role of the IHE in the MSP theory of action is based on the belief that partnership between K-12 and IHEs is mutually supportive and can enhance learning, cultural awareness, and teaching practices for both partners.

---

[6] Math & Science Collaborative Journal, v11, p.20.

[7] Network Connections is a semi-annual gathering of LATs from across the SW PA MSP together with appointed teams from other non-participating districts in the region to consider current research and materials to impact K-12 science and mathematics education.

Relationship Between Logic Model and Evaluation Questions

The MSP logic model and research evaluation questions form a framework to guide the evaluation. The first research evaluation question addresses the need to provide formative assessment and documentation of MSP activities and corresponds primarily to the Intervention Strategies and MSP Activities listed in the logic model. However, the discussion of MSP activities also has implications for the Inputs and Outputs. The successful implementation of the MSP activities depends on the resources available to generate these activities. Moreover, successful implementation of MSP activities is reflected in the Outputs that are generated. For the MSP, this is evidenced by the quantity and quality of instructional leaders, curriculum materials, and opportunities for interactions. The second and third research evaluation questions assess some of the key outcomes of the MSP — changes in institutional practices, changes in support structures at both the district and IHE levels, and changes in instructional practices at the classroom level. Although these questions primarily track the Mid-term outcomes in the logic model, the Short-term outcomes are important precursors that indicate the likelihood that these Mid-term outcomes will be achieved. Finally, the fourth research evaluation question focuses on the bottom line of student outcomes and changes in course taking practices. In essence, this question is asking whether the project achieved its goals, which are outlined by the long-term outcomes of the logic model. One of the most important aspects of this assessment will be to determine whether achievement of these long-term outcomes can be linked to MSP activities.

Limitations of the Study

This evaluation is designed to be selective in its data collection and analyses, primarily assessing the project's achievement of its goals and the major pathways toward achieving those goals. In addition, this study is not being implemented as a randomized experiment with a control group, and this fact will limit the ability to make definitive causal claims about the MSP's impact. Another important consideration of this study is the lack of a statewide science test. Although the project is administering a science exam to students in participating districts, the exam is not administered widely outside the project, so there are few options for comparing the science performance of MSP districts to external references such as comparison districts. Until recently, the state administered a math test to students in only a few grade levels, so scores from immediately prior to the start of the MSP are not uniformly available for all students to serve as baselines for math achievement. At the conclusion of this five-year evaluation, lags in reporting of achievement data may force the evaluation to confine its analyses to test results through the 2006-07 academic year, covering only three full years of MSP implementation in the school districts. If the MSP intervention strategies require more than three years to affect achievement outcomes, the evaluation may not be able to detect this impact. To supplement and illuminate the quantitative analyses, the project relies heavily on qualitative self-report data, through surveys, interviews, and focus groups. However, this data is subject to common biases associated with qualitative data, for example, if respondents answer survey or interview questions inaccurately or if the sampled districts for case studies are not representative. The analyses discussed in this Year Four report focus on a very narrow subset of data and the preliminary findings reported herein cannot be extrapolated across the entire data set. Additional details about the limitations of specific analyses are described in context throughout this report.

Organization of this Report

This report differs from earlier interim reports in that it uses participation levels to focus on a narrow subset of data from MSP partners. Chapter 2 describes the rationale for choosing this focus and describes how the evaluation team measured participation. Chapter 3 presents findings from the K-12 Case Studies. Chapter 4 presents the results from the principal survey. Chapter 5 describes the various types of student achievement data available for this evaluation, and presents two statistical models testing the relationship between measures of participation and student achievement outcomes. Chapter 6 presents preliminary findings from the IHE Case Studies. Finally, findings from Chapters 3 through 6 are summarized briefly in Chapter 7, along with several issues arising from these findings for the project to consider as it moves forward.

2. Measuring Participation

The MSP logic model presented in Chapter 1 indicates that the intervention strategies are expected to lead to mid-term outcomes, which are, in turn, expected to lead to long-term outcomes. In other words, *participation* in the project's activities leads to *implementation* of reformed policies and practices, which leads to *attainment* of the project's goals. Although the bulk of evaluation data collection efforts focus on measuring outcomes, participation is a crucial first step in the process leading to attainment of project goals. This chapter discusses the available measures of participation, and describes one selected to serve as an organizing theme for this year's evaluation report.

Rationale for Measuring Participation

When a school or school district opts to join a major reform effort such as the MSP, it is not a foregone conclusion that the reform will be implemented successfully or achieve the desired outcomes. In fact, research on major school reform efforts finds that few schools adopting such reforms fully implement the program designs, and that educational practices in schools adopting reforms closely resemble those in schools not adopting them (Vernez et al., 2006). Even in schools where reform-oriented instructional practices are implemented in classrooms, producing measurable effects on achievement is not a certainty. For example, a study of reform-oriented teaching practices in math and science found nonsignificant or weak positive relationships between those practices and students' scores on state standardized tests (Le et al., 2006). Thus, this evaluation confronts a challenge faced by many prior evaluations of major school reform efforts: effects on student achievement are difficult to measure and attribute to the reform. Fortunately, the evaluation does not focus exclusively on attainment of the project's goals, but includes measures of outcomes along the path from project activities to the goals.

As outlined in Chapter 1, the MSP intervention strategies include activities that: provide professional development for content and leadership, support the alignment of curriculum and refinement of courses and pedagogy, and disseminate and support the use of research-based resources and tools. Effects of the MSP on both mid- and long-term outcomes are likely to be stronger in districts that more fully participate in these activities, and weaker in districts that participate to a lesser extent. Measuring participation is therefore a key element of this evaluation for two reasons. First, participation may help to explain differences in observed effects across districts. Second, evidence of a correlation between participation measures and changes in outcomes can help to support a plausible argument that the MSP is responsible for those changes.

Measures of K-12 Participation

The MSP project maintains a comprehensive participation database of information on every educator in MSP school districts who is eligible to participate in project activities, as well as on other participants who are members of non-MSP school districts, IHEs, or other organizations. The database includes the number of hours the individual engaged in each MSP activity, the dates of those activities, and other information about the individual such as institutional affiliation (district, school, IHE, etc.), job function (teacher,

principal, superintendent, guidance counselor, professor, etc.), subject areas taught (math, science, special education, etc.), years of experience, and demographic information such as race/ethnicity and gender.

The evaluation team used the information in this database to develop a measure of district participation in MSP activities over the first three years of the project, based on a January 2007 snapshot of the database. The snapshot contains participation records for 12,100 individuals. These records contain complete information on all MSP activities held during the first three years of the project, as well as some information from the fourth year of the project. The Year Four information is not included in the current analyses because it is incomplete and not uniformly reported across districts.

As an initial step in creating a participation measure, the evaluation team totaled the participation hours for each individual (educator/administrator, etc.) across all activities over the three years, and aggregated these totals to the school district level. Focusing on the 48 NSF and PDE school districts involved in the MSP since inception, the result is a single value for each district representing the total number of hours of participation by all district staff. This total ranged from 122 to 5,259 hours, with a median of 2551 hours. This range indicates large variation in total participation hours among the MSP districts.

Because there is also considerable variation in the sizes of MSP districts, counting total hours is an imperfect measure for comparing district levels of participation. Across the MSP districts, student enrollment ranges from 687 to 7,646, and teaching staff size ranges from 54 to 512 (U.S. Department of Education, 2006). A district with a small staff that participates to the full extent might not accumulate as many hours as a large district that participated to a lesser extent. For this reason, the evaluation team created an adjusted measure of participation that divides each district's total hours by its total teaching staff size. This measure, *total hours adjusted for staff size (THASS)*, ranges from 1.37 to 45.11 across the 48 MSP districts.

Limitations of the K-12 Participation Measure

For the analyses in this report, the evaluation team selected THASS as a measure of participation, rather than unadjusted total hours of participation. This is not to say that THASS is a perfect measure, and for future analyses the evaluation team will seek better measures of participation. For example, THASS does not account for specific educator roles that may be more influential in translating MSP participation into the implementation of policies and practices (such as the roles of district or school administrators or leadership action team members). Moreover, THASS does not differentiate between MSP activities that may be more or less effective at producing the desired outcomes. A more refined measure of participation might account for differences in the importance of specific activities or participant roles.

There are additional limitations of THASS. It may underestimate participation in large districts because many MSP roles and activities have limited capacity. That is, opportunities to accumulate participation hours do not scale linearly with district size. In preparation for the analyses that will appear in the final evaluation report, the evaluation team intends to work with the PI to develop a more refined participation measure that more accurately represents the extent to which each district has taken full advantage of its opportunities to participate.

Finally, THASS is necessarily limited as it only provides information on participation hours. It does not account for participants' level of engagement in MSP activities or buy-in to the MSP intervention model. Two districts with equal participation hours might differ extensively on the quality of implementation if, for example, participants in one district are very engaged in MSP activities and enthusiastic about the MSP model and the participants in the other district are less engaged or perhaps more apathetic about the MSP model. Information from the K-12 case studies may help to understand the variation in quality of implementation within specific districts, but resources do not enable the evaluation team to systematically determine the quality of implementation across all districts, which would require extensive observations.

Assessing Participants' Alignment with Project Goals and Philosophy

In March 2007, MSP staff developed a systematic method for the PI, project directors, and coordinators to assess the extent to which each participant has aligned with the goals and philosophy of the project. The staff created and used a rubric, titled "Readiness for the Journey," with five levels: "*in synch" leads*, *endorses*, *listens*, *going through the motions*, and *actively resisting*. Each level has a set of descriptors that help to define these terms. For example, the descriptors for "*in synch"* include: owns the goals; knows how to implement the philosophy; coordinates connections between the MSP and other district initiatives; "talks the talk, walks the walk," and leads others; enthusiastically implements, even beyond what is required; and problem solves to overcome barriers. At the other end of the spectrum, the descriptors for *actively resisting change* include: would like to avoid changing their working routines; may arrive late, leave early; inattentive (off-task) behavior during sessions; and may feel implementation is not in their best interest. It is worth noting that these descriptors contain indicators of both depth of participation as well as willingness and capacity to implement.

Using this rubric, the staff members classified each participant with whom they had direct contact through March 2007. Participants with whom staff did not have contact were not classified. Where more than one staff member classified a participant, the scores were averaged. In this report, these scores are referred to as the *readiness scores*. Even though a systematic process was used to create the readiness scores, they are somewhat subjective in that they are based on staff interactions with participants and their perceptions of the individuals.

To the extent that these readiness scores capture information on the quality of participation and implementation, they hold promise as a complement to simpler measures based on participation hours. The readiness scores were aggregated to the district level, by averaging those of each district participant who has one. The result was a readiness score for each MSP district, ranging from 1.0 to 4.1. As part of an exploratory analysis, the evaluation team then examined the relationship between the readiness scores and THASS. As shown in Figure 2-1, the district-level scores are not correlated with THASS, supporting the notion that they measure an aspect of participation that is distinct from time-based measures. In future months the evaluation team will explore the relationship between the readiness scores and time-based participation measures, and whether a combination of the two might be a useful predictor of outcomes.

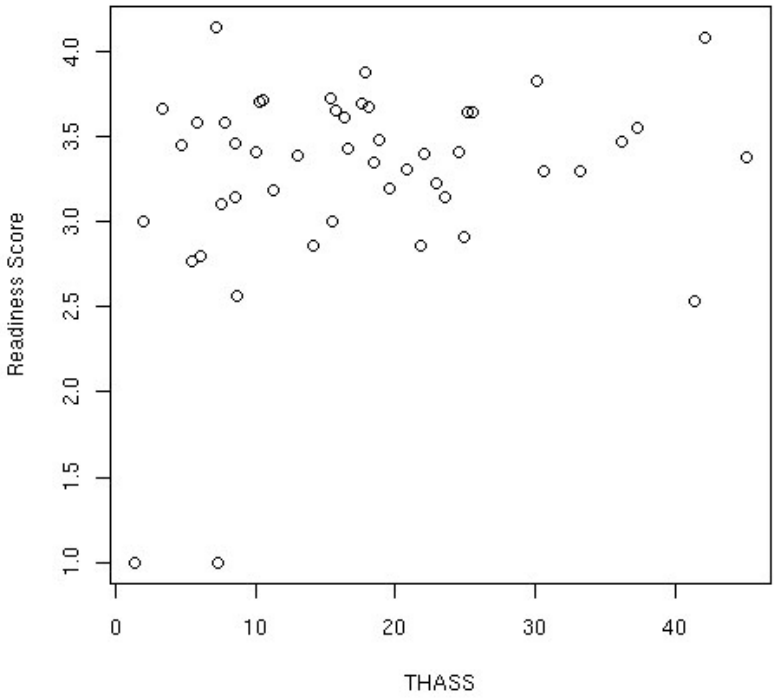


Figure 2-1: The readiness score developed by MSP staff is not highly correlated with THASS, suggesting the two variables may measure distinct aspects of participation.

IHE Measure of Participation

To measure participation in IHEs, the evaluation team utilized the NSF Management Information System (MIS) database, which contains self-reported information on the number of hours faculty members spend participating in MSP activities. Because of the limited number of IHEs in the project, the evaluation team considered participation at the individual level to be a better unit of analysis than the institutional level. Fifty-nine faculty members from the four IHEs are included in the MIS database. In the MIS, faculty members annually estimate the time they spent participating in MSP activities in the prior year, by selecting one of the following responses: 0-20 hours, 20-40 hours, 41-80 hours, 81-160 hours, 161-200 hours, or more than 200 hours. To create a total estimate of participation, a single value was assigned to each response option by using the midpoint of each range (i.e., 10, 30, 60.5, 120.5, and 180.5) and 200 for the more than 200 hours response option. Each faculty member's participation was aggregated over the three years by summing the yearly values. The result ranged from 10 to 600 hours per faculty member.

Report Focus: The Relationship Between Participation and Outcomes

For this year's annual report, the evaluation team used participation levels as an analytical focus to identify promising practices among MSP partners. For K-12 districts, the evaluation team opted to use the

THASS measure of participation to contrast high- versus low-participating districts. To determine the range of values that comprise both the high and low ends of district participation, the THASS values were plotted on a logarithmic scale. The team observed breakpoints near the high and low ends of the range (see Figure 2-2). Using these breakpoints as a cutoff, eight districts were selected as high-participating districts, with THASS values ranging from 30.2 to 45.1 (median 36.8), and seven districts were selected as low-participating districts, with THASS values ranging from 1.4 to 6.1 (median 4.7).

For IHEs, the evaluation team identified faculty members who were high-participating, based on the estimated total hours of participation derived from the MIS. Information obtained in prior-year IHE interviews was also considered. Faculty members with more than 200 total hours of participation and who were actively involved for more than one year were classified as high participating (n=21). In three cases, faculty members with more than 200 hours of participation were not classified as high participating because they had left their IHE or they had become very involved only in Year 3.

Using this participation theme, the following chapters examine the low- and high-participating MSP districts and high-participating IHE faculty exploring factors that influence participation and the effects of participation on mid-term and long-term outcomes. Where appropriate, the discussions are supplemented with participation measures other than THASS. For example, in Chapter 3, the readiness score is also used because it represents another facet of participation that is empirically different than THASS, and in Chapter 4, principals' participation in *Lenses on Learning* is included as a more focused measure of principal involvement in the MSP than the THASS measure.

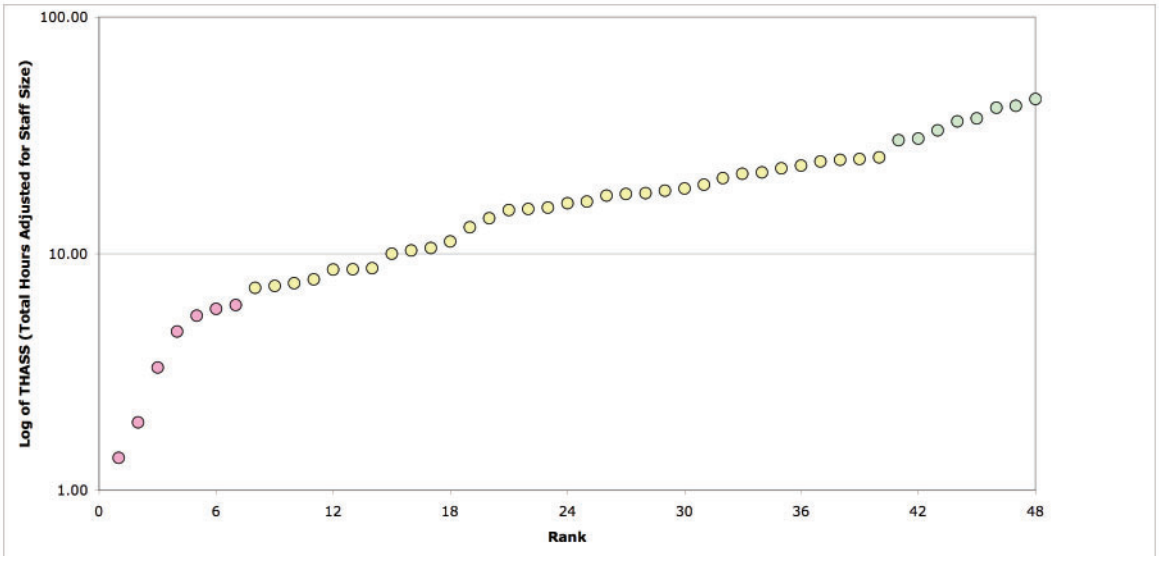


Figure 2-2: Distribution of MSP districts ranked by THASS. Seven districts at the bottom of this distribution were selected to represent low-participating districts, and eight districts at the top were selected to represent high-participating districts.

3. School District Case Studies

The K-12 Case studies are designed to add a depth of contextual understanding to the other data sources. By capturing a detailed picture of the MSP from the school districts' perspectives, the evaluation team can provide a more comprehensive explanation of the findings from achievement data, teacher and principal surveys, and interviews of MSP staff and IHE faculty. The evaluation includes 12 K-12 case studies, seven school districts from the NSF-sponsored group and five from those supported through the PDE.

An important goal of the case studies is to provide a detailed description of districts that faithfully implement the MSP intervention strategies. This description may help explain the reasons why some districts do not implement with fidelity. Thus, demographics and early measures of participation and implementation were used to identify a representative set of districts for inclusion in the case studies. The case study districts were selected at the beginning of the project, in late in 2004 and have been reviewed annually to ensure that they remain representative of the MSP-participating districts. This chapter discusses data collected from four high-participating districts (herein will be identified as H1, H2, H3, H4) and three low-participating districts (L1, L2, L3), as determined by the THASS participation measure described in Chapter Two. In the third year of the project, data collection for the case studies primarily focused on districts at the highest level of participation. As a result, comparing high- and low-participating districts may be limited by lack of information from the low-participating districts. In such cases, information on high-participating districts is included even if there is not comparable information from low-participating districts.

Case Study Data Collection

In Year Four as in previous years, the K-12 case studies included observations of the summer Teacher Leader Academies and follow-up sessions, observations of on-site academies, observations of K-12 classroom instruction, and interviews with school, district, and IU personnel and project staff. However, a focus in the case studies on teachers with "high level participation" at Leadership Action Academies (LAA), Teacher Leader Academies (TLA), and other MSP activities, resulted in more classroom observations across fewer teachers in Year Four. For low-participating districts, observations of at least one of their on-site academies occurred if it was provided in their district, and follow-ups to last year's interviews were conducted. Detailed information on the observation protocols that were used can be found in prior reports (Williams et al., 2005; Williams et al., 2006). Over the course of the three years of data gathering, revisions were made to the interview protocols in order to address issues raised in previous years. Copies of the interview and observation protocols are available upon request.

The MSP staff retreat (August 2007) and LAAs were important components of the observation data because they demonstrate how the MSP staff review and share district information and how their understanding of the districts transfers into the development of the LAAs. In addition to observation data, the evaluation team examined district files maintained by the MSP to document attendance and communication with project staff.

Participation in LAAs and TLAs

The LAAs help to set the tone, plans, and goals for district action. During Day 1 (of 4) of this year's LAAs, the Leadership Action Teams (LATs) performed a "health status" exercise as an extension of their annual review and self-reflection. Participants compiled a list of characteristics that describe a "healthy" district in terms of MSP participation and subsequent performance and capacity to implement reform. With that list in mind, the teams were asked to categorize their district health status as "superior, excellent, good, poor, or intensive care". Based on this categorization, districts then reviewed their district achievement data, reflected on the past year's activity and accomplishments, and began the work of updating their district action plan (among other documents) for the upcoming year, adjusting for emerging issues from the health status assessment.

The districts identified as high-participating (H1, H2, H3, H4), recorded high attendance across all four LAA sessions, and lead administrator attendance in three of the four districts. Feedback on these academies from members of high participating LATs noted the MSP's "contribution to helping a district create a conceptual and practical framework to organize human, technological and curricular resources to support students and faculty successes". LAT members highlighted the importance of having the time to create long-term goals for themselves. These districts' lead administrators shared the challenge of keeping a central focus in their district while weaving together different reform initiatives. One district Superintendent reported that a strength of the district's involvement in the MSP has been the opportunity to "integrate our MSP initiatives and activities into related district changes (e.g. our curriculum revision cycle sessions; our new K-12 curriculum committees)".

In contrast, the low-participating districts have had inconsistent or no participation at the LAAs. In an interview with three teachers from a LAT at a low participating district, the teachers recognized the challenge of having a new superintendent and no lead administrator on the LAT this past year to help guide their efforts with the project, but remained hopeful that with time the new superintendent would help make changes in the district. It is worth noting that although this superintendent did not participate in MSP activities in the first part of Year Four, this changed in the second half of the year. This district is moving toward greater participation in overall MSP activities such as LAAs and High School Science Summer TLAs and follow-ups in Year Four. Their challenge remains having a small and veteran mathematics department and a continued struggle to identify teacher leaders within the district to support math reform. The district leadership remains committed to making changes and has continued to seek support from their MSP staff contact.

The MSP staff has made efforts to engage L1 through continued outreach efforts and correspondence. Nonetheless, the district has remained uninvolved. While some teachers indicated interest in MSP offerings during Year Three, key project personnel have not responded to the MSP's outreach. Knowing how quickly a district's status can change, the MSP stands ready with an "open door policy" should the district renew its commitment to participate. L2 district experienced major changes in its organization. As a result of its poor performance under NCLB and continuing financial distress, the district has been reconstituted and is being managed by another non-MSP district. It appears unlikely that this district will re-engage as an MSP participant.

Findings from the TLAs

Each project year's TLAs begin at the close of the project year in the summer months. Follow-up sessions begin in the following project year as the school year begins. Follow-up sessions enable teacher leaders to address specific challenges they encountered implementing the on-sites academies in their home schools. As a result, the Year Four report includes findings from the Year Two TLA follow-up sessions held during the 2005-06 school year, the Year Three Summer TLAs, and the first two of four follow-up sessions in Year Three.

Among the high-participating case study districts, attendance at Summer TLAs and the first two follow-ups[8] remained high. The experiences of these districts highlight the role of school and district leadership in how reform initiatives are perceived and implemented. Leadership and teacher enthusiasm appear to both be mutually supportive; both are necessary for creating conditions for implementing reforms.

For example, H4 appointed teachers to each of the new TLAs, participated in summer academies and follow-ups and initiated the on-site professional development. H3 has teachers participating in two of the new TLAs and two other high-participating districts have teachers participating in one. Early reports of on-site activity for these academies[9] also show implementation within the districts.

Implementation of On-Site Academies

Districts have indicated that a major challenge for any reform initiative is finding ways to provide time for necessary professional development. This issue has emerged in the MSP as well. Many districts are unaccustomed to devoting focused professional development for a specific reform across five years. As such, districts must adapt their structures and practices to accommodate reform and the reform (in this case, the MSP) must also adapt to district context. Two high-participating case study districts provide evidence of ways they adapted MSP professional development guidelines to their district context.

H4 noted "these past two years have been a challenge in finding time. Now as we move to lesson study, we juggle our professional development into our morning meetings, monthly in-service, voluntary after-school time and occasionally use substitute teacher coverage. Teachers may get 45 minutes here and 20 minutes there". This district's LAT point person opened an in-service lecture stating, "We can't do Lesson Study the MSP way. We have to do it the [H4] way". A follow-up interview with H4's point person further explained that district context "mandates that we be creative in how we get the information across". Instead of creating 5-8 person study groups to engage in lesson study, this district has implemented lesson study using the Thinking Through the Lesson Protocol (TTLP) and 5-E instructional model tool as a ways for individual teachers to reflect on and revise their individual lessons under the guidance of MSP trained district teacher leaders. They do this during an hour of in-service as reflective practice by adapting existing lessons to incorporate features of inquiry.

---

[8] As mentioned previously, the reporting cycle covers only the first two of five follow-ups.

[9] As reflected in 1/18/07 MSP Master Spreadsheet Data exported from the MSP database.

By contrast, H2 has two lesson study groups that began in September 2006. Each group has six members from across disciplines including mathematics, science, history and technology. This district is requiring that all of its teachers engage in one of two different professional development experiences. As a result, teachers of subjects other than math and science have joined lesson study groups. Each team has completed nearly 20 hours of work on planning their lesson. This time does not yet include time for teaching the lessons, debriefing, and revising. Teachers indicate they are doing the lesson study work outside of their school day. One teacher said, "the collaboration is more meaningful than the lesson--hearing the different perspectives regarding our various pedagogies and content was so worthwhile." However, this teacher also added that the lesson study portfolio is inefficient, difficult, and intense to complete at the end of an extended work day. Another teacher noted, "maybe [just using] the tuning protocol (TTLP) would work better for us in terms of time" as opposed to following the Lesson Study Portfolio which includes more planning items such as establishing group norms, defining long-term goals, determining content area, unit and lesson affiliated with the long-term goal, debriefing and revising items in addition to the lesson protocol.

The MSP recognizes that adaptation is linked to both integrity of the intervention and, ultimately, sustainability. The MSP recognizes that some degree of adaptation is necessary in order for the intervention strategies to be sustained. Thus, the MSP is attempting to clarify a continuum of adaptation for teacher leaders to consider where the integrity of the intervention is not compromised. Because the on-site academies are a critical element to developing communities of learners, the evaluation team will continue to monitor, through the case study districts, the types of modifications that are being made to the on-site academies, the conditions that induce them, and the impacts of the changes.

Successful implementation of the MSP project may begin with successful training of TLAs and implementation of on-site academies, but actual change in policies and practices indicate a deeper realization of MSP goals. There are a number of changes observed in high participating case study districts that are aligned with the short- and mid-term outcomes in the logic model and lend evidence for the notion of deeper impact. They are discussed more fully in the next section.

Use of Data in Decision Making

The MSP encourages districts to analyze their school performance data and use it to develop district plans. An interview with a leadership action team member at H4 indicated that the "data analysis that occurs each year with the MSP has helped to both guide planning and has influenced decision making" with regards to resource allocation, professional development planning, curriculum choices and the like. Participant responses from LAA Day 2 also indicated that "data driven decision making has been enhanced with some of the analysis pieces that the MSP has provided" while another participant noted "the MSP has underscored the importance of examining and understanding data to assess student progress". Use of data in decision making is a common and seemingly frequent response among participants as a major change in district policies and practices. A teacher leader in H3 said, "we do more curriculum mapping [as a result of MSP participation]…definitely the MSP has opened the eyes of some administrators to use data to drive their decisions".

Collaboration Across Schools and Districts

At the K-12 level, high-participating case study data indicate that there has been an increase in communication across schools within a district, and even across districts. One H3 Teacher Leader noted she has benefited greatly from the collaboration that has occurred with other teachers from other districts. She reported she has called teachers that she met from other districts through the MSP to seek advice regarding on-site academy facilitation and even individual student concerns. Evaluation responses from TLAs further highlights the benefits. Comments ranged from, "being among a group of elementary teachers who "teach math" gave me exposure to all levels of understanding and perspectives" and "learned a lot from districts sharing their math needs and course of action concerning professional development. Ideas shared today will help me facilitate school improvement plans in my district". Feedback from the Fall 2006 Network Connections conference and the Superintendents luncheon during Day 2 of the LAA also noted increases in teacher collaboration within and across districts as a benefit of involvement with the MSP.

Awareness and Knowledge of Research-Based Instructional Practices and Materials

The MSP's professional development activities are essential to ensuring that teachers not only understand *how* to change instructional practices, but also *why* such changes are important. Awareness of the larger purposes behind these changes is a critical component of sustainability as it provides teachers with a foundation that can last beyond the funding period of the project. Another participant said, "the MSP has contributed to our success by reinforcing pedagogy regarding best practices in the teaching of math and science, including the basic principles of authentic instruction, inquiry-based learning and constructive knowledge". Participants from Teacher Leader Academies often comment directly on their understanding of research-based instruction and materials. One participant commented, "I have realized that you have to know where you want the students to go with the lesson so you can scaffold them rather than confuse them". Another participant commented, "understanding how the child solves the problem allows me to help solve misconceptions the child has…and has provided me with a first step toward making changes". Further interview data from high-participating district administrators and teachers have suggested this awareness and understanding of research-based instructional practices and materials has impacted curriculum choices. The point person for H4 noted, "because of our partnership [with the MSP] our district is successfully transitioning to a program that incorporates high levels of hands-on activities".

Increasing Teacher Content Knowledge

As outlined in the MSP theory of action, increasing teacher content knowledge is important to achieving changes in teacher practices. Content deepening is intended to occur not just in the Content Deepening Seminars but in all teacher leader academies and educator networks. Increased content knowledge makes teachers more comfortable with using many of the strategies advocated by the MSP. Participants are frequently asked in their evaluations of various MSP activities, "To what extent did looking at the mathematics or science in the lesson deepen your understanding of the content of the unit?" Evaluations

show the mathematics and science educator network sessions were rated highly. Many respondents indicated a greater understanding of particular concepts. A participant from an Early Learners Teacher Academy commented, "I will now view the 'Big Ideas' math concepts in a different way. Classifying, patterning, sorting, and using manipulatives during counting and problem solving have meaning and importance to help children understand and develop these concepts".

Classroom Instructional Practices in Math

Data collected from Year Four event observations indicate that math teachers are benefiting from MSP activities that focus on the use of active learning, and are pleased with what they are learning. Moreover, in all but one of the case study classroom observations in high-participating districts, elementary math instruction was rated as *effective* to *exemplary* and including active student learning opportunities. Generally, classroom observations in districts with less participation yielded somewhat lower ratings. Instruction was focused on the process as well as the outcome with teachers posing questions such as "How did you get that?" and "Who can help [student's name]?". Class activities were purposeful and engaging. A third-grade math class of 17 students from H3 exemplifies these lessons. The students had time to work individually, within small groups and as a large group. The instruction was well-thought out and the instruction was seamless between activities. The teacher noted that he has 30 years of experience and has found that since his school has departmentalized he has been able to focus all of his energy on improving his math instruction. It was apparent from the students' engagement and eagerness that the instruction had caught their attention. This teacher participated in MSP Teacher Leader Academies and noted that his experience with the MSP has supported his professional development.

Observations from secondary mathematics classroom practice indicated a wide diversity of instruction across high-participating districts. Some instruction exemplifies MSP teaching reform while others reflected more traditional rote methods of instruction. Classroom instructional practices varied across individual teachers within each high-participating district and even across classes taught by the same teacher. Interviews with these teachers explain a lack of time to be a challenge to reforming their practice. A secondary math teacher and district teacher leader at H3 commented, "I don't have the time to change all of my lessons at once. I have been able to add more inquiry to some of my lessons more easily than others, especially in my geometry classes. On a whole, I do know that I ask more questions [of my students] rather than provide them with all the answers but as far as refining my lessons, I do what I can in the summers." This teacher has taught for six years and teaches six of nine periods a day including algebra, geometry, and calculus.

Some factors seemed to limit high quality learning for students. Interruptions to classroom instruction, noted across many observations, seem to jeopardize opportunities for learning. A Grade 5 mathematics classroom with 25 students proved to be challenging for the classroom teacher. The students were engaged in a hands-on activity in small groups using a local restaurant's menu to solve problems. Students were coming in and out of the classroom because of an external need to collect fundraising items from the main office. This appeared to distract both the teacher and students from the lesson. The teacher found it to be challenging to move about to each group and provide support to keep the students on task while they were being called to the office. This teacher further noted that while the MSP activities were

"helpful" he has found them to be less useful in "his district context". His class is a mix of rural and urban students and he finds it challenging to get all of them onto the same page.

Classroom Instructional Practices in Science

Through the science curriculum framework and the high school science TLAs, teachers are encouraged to include inquiry when teaching scientific materials rather than relying on lecture alone. The MSP advocates for student engagement with hands-on, interactive activities that challenge students to think and act like scientists.

Across the high-participating districts, case study observations tend to rate high school science classroom instruction lower than math. While instruction was typically purposeful and well-planned, implementation of the lessons seemed more challenging. Lessons were often "short circuited" by a lack of time for lesson debrief or because the teacher did not have enough time to move from group to group as s/he was often caught up in trying to "fix experiment mishaps" among the groups. As noted in the Year Three evaluation report, science teachers report challenges to enacting their lessons, such as, a lack of time to implement a lesson fully due to short class periods, and a loss of continuity in labs because they must be held over several days. Through teacher interviews, the evaluation team learned that teachers have appreciated and benefited from the TLAs and on-site academies. One high school Biology teacher noted, "Being a scientist, I know what good science looks like. I know that experimentation and questioning works. It's the lack of time that is challenging. I used to have planning periods and since other teachers complained, administration took it [sic] away. So I just do what I can".

Challenges in Low-Participating Districts

Case study findings indicate that low-participating districts are among the most economically and educationally challenged districts in the MSP. These districts are often inundated with reform efforts focused on improving student achievement, yet are unable to meet the requirements for reform because of limited capacity in human capital, funding resources, and infrastructure as well as competing reforms that diffuse the focus of any one initiative, including the MSP. The result is often unfocused activity that can create chaos in an already stressed environment. The added pressure of high stakes accountability (including threatened and actual state takeovers) and public scrutiny by stakeholder groups can make focused and sustainable reform particularly difficult. Among the case study districts, those that fall into the low-participating category have all reported difficulties in implementing reforms through the MSP. While the case studies have documented this phenomenon, further exploration of the underlying issues is a crucial research agenda beyond the scope of this evaluation, though some discussion of these issues occurs in Chapter 7.

4. Principal Survey

The evaluation team administered a principal survey in 2004 to collect baseline data on principals' views, attitudes, policies, and practices on a range of subjects related to the MSP. Prior evaluation reports discussed the results from this baseline principal survey (Williams et al., 2005; Williams et al., 2006). In 2006, the evaluation team administered the same survey to assess principals' responses after further implementation of the MSP project. This chapter discusses changes observed from baseline and examines the relationship between participation and survey results.

Survey Response and Principal Demographics

The evaluation team surveyed one principal in each of the schools in the original 48 MSP districts. The response rate was 71% (142 responses from 201 principals sampled) for the 2004 survey and 92% (181 out of 197) for the 2006 survey. Surveys were received both years from 129 schools. From 2004 to 2006 there was considerable turnover among principals in the MSP schools. As a result, the 2006 sample included only 118 of the 201 principals who had been sampled in 2004, and 84 of these responded to both surveys. While the survey can measure change in school-level practices and policies even with this turnover, attributing any observed changes to the MSP intervention will be problematic because the new staff might have brought those changes even in the absence of the MSP. Moreover, the turnover is a concern because it might diminish the effect of the MSP if project-induced changes are not sustained in a school when a participating principal departs.

Notably, the turnover of principals resulted in a shift in some of the demographic variables summarized in Table 4-1. In general, principals responding in 2006 reported fewer years of experience in their district than those who responded in 2004 (Cochran-Armitage Trend Test $p<0.01$). In 2006, 9% of principals reported having more than 10 years of experience as a principal in the district while 19% reported having one year or less of such experience. In 2004, these values were approximately reversed: 19% reported greater than 10 years of experience while 11% reported one year or less of experience. A similar trend, though non-significant, appears in the educational degrees reported, with fewer principals in 2006 holding a doctorate degree and more with a master's degree plus additional course work.

Table 4-1: Principal Survey Respondent Demographics

Characteristic		2004 Survey (%)	2006 Survey (%)
Gender	Female	46	47
Race	White	90	91
	Black	6	7
Years as Principal in the District*	>10	19	9
	7-10	24	20
	4-6	26	29
	2-3	19	24
	0-1	11	19
Highest Degree	Doctorate	28	21
	Masters +	29	35
	Masters	42	43

* $p<0.01$ (Cochran-Armitage Trend Test)

Methods Used to Account for Survey Non-Response

As mentioned above, the primary intent of the principal survey is to gain school-level measures of principal views, attitudes, practices and policies and to measure change from baseline to post-implementation. In order to make valid inferences across all of the MSP schools, missing values for survey non-responders and item-level non-response by survey responders are imputed. This enables inferences about the full population of schools, rather than forcing attention to be restricted to only those from which survey responses were received. Multiple (5) imputations of missing values were performed for the item-level responses using a multivariate technique consisting of a sequence of stepwise regression models. Imputations were done separately for each survey year. Possible predictor variables in each model consisted of all item level responses for a given year, as well as principal, school, and district demographics. Because of high survey response rates, the population estimates are relatively precise. All subsequent analyses adjust for the between- and within-imputation variance; that is, all subsequent statistical tests account for the uncertainty introduced by the imputation.

Survey Content and Results

The principal surveys include items adapted from survey instruments developed by Horizon Research, Inc., the Center for the Study of Teaching and Policy, and the Center for Research on the Context of Teaching; and principal rubrics developed by Richard Halverson at the Wisconsin Center for Education Research. It was designed to capture changes in principals' views and attitudes across 6 areas:

- Views and influences on math instruction
- Views and influences on science instruction

- Practices and policies regarding curriculum, instruction, assessment, and professional development
- District and IU support for improving schools
- Impact of the MSP project
- Factors and incentives affecting MSP participation (2006 survey only)

The Year 3 report described the factor analysis used to identify survey scales, which represented subsets of highly correlated survey items (Williams et al., 2006). These scales were linked to short- and mid-term outcomes in the logic model to help determine how survey results aligned with project outcomes.

Some of the principals' responses on the survey scales changed between 2004 and 2006. The most significant changes were found on scales related to leadership skills and district and school level practices and policies. Tables 4-2 and 4-3 indicate changes in principal responses that are significant. The titles of the tables refers to the short- or mid-term outcome that link the scales. The tables include the scale description in bold and the survey items that are included in the scale below the bold heading. Statistical significance for the bulleted statements about each scale is reported as a p value. In some cases, significant changes were observed for science scales but not mathematics. For some of the scales, the mathematics scores were higher than the corresponding science scales in the 2004 survey. If those scores were much higher, it might not have been possible to measure a significant change due to a "ceiling effect". However, this does not appear to be the case, as three of the four highest mean scale scores in 2004 showed significant increases in 2006. In the following discussion, effect sizes that are all based on responses to survey scales reported on five-point Likert scales are reported.

Principals' responses to three scales associated with leadership skills changed significantly. In the 2006 survey, principals responded higher on the five-point Likert scale "strongly disagree" to "strongly agree" to being knowledgeable about national standards for science, being knowledgeable about the district science curriculum, and being prepared to support teachers in implementing the science curriculum. They also reported an increase on the parallel mathematics questions, although the difference was not quite significant. On a scale of "not comfortable" to "very comfortable", they reported being more comfortable serving as math and science instructional leaders, including discussing concrete examples of instructional practices with teachers, observing classrooms, examining student work, and providing feedback on teaching. Finally, principals responded higher on a scale of "inhibits effective instruction" to "encourages effective instruction" that developing a professional learning community—giving teachers time to prepare lessons, work with other teachers, attend professional development—encourages effective instruction. However, the effect sizes are fairly small (0.1-0.2) for all of these except the last scale (effect size for science of 0.4).

Principals' responses concerning district and school level policies and practices also changed significantly. On two scales, which include items such as the regional curriculum framework, testing policies, access to computers and calculators, and the quality of district adopted instructional materials, principals reported higher on the scale that these items "encourage effective instruction". The effect size for these changes is greater than those associated with leadership skills. For the mathematics questions, the effect sizes are 0.2 and for science, 0.5.

Table 4-2: Administrative Leadership Skills

Scale: Principal Knowledge of and Support for Implementing Curriculum and Standards

Example of Survey Items:
- I am knowledgeable about current national standards
- I am knowledgeable about the curriculum in my district
- I feel well prepared to support teachers in my school in implementing the curriculum

2007 Findings:
- Principals reported higher values of "agreement" in 2006 than in 2004
  Math $p=0.08$; Science $p<0.05$
- Lenses on Learning participants reported higher values on the mathematics scale than non LoL participants
  Math $p<0.05$

Scale: Principal Support for Teachers To Develop a Professional Learning Community

Example Survey Items:
- Time available for teachers to plan and prepare lessons
- Time available for teachers to work with other teachers
- Time available for teacher professional development

2007 Findings:
- Principals reported higher values of "encourages effective instruction" in 2006 than in 2004
  Science $p<0.0001$
- LoL participants reported higher values on this scale than non LoL participants
  Math $p<0.01$; Science $p<0.05$

Scale: Principal Comfort In Serving as a Math or Science Instructional Leader

Example of Survey Items:
- Discussing concrete examples of instructional practice with teachers
- Holding pre-observation conferences with teachers
- Observing classroom instruction and examining student work
- Providing teachers with feedback on teaching

2007 Findings:
- Principals reported higher values on "comfortable" scale in 2006 than in 2004

The survey also included questions about the impact of the MSP on instructional practices overall and the impact of individual MSP activities on instructional practices. In 2006, principals reported that the MSP had a greater overall influence on mathematics and science instructional practices than in 2004. Their responses were 0.2 points higher ($p<0.05$) on the five-point Likert scale ("not at all" to "a great deal"). The percentage of principals reporting that the MSP has a "great deal" of influence increased from 18 to 30 percent (Figure 4-1). However, when focusing on specific MSP activities rather than the MSP overall, principals did not report significant increases in the influence on mathematics and science instruction.

Table 4-3: Changes in District and School Level Policies and Practices

Scale: Effect and Availability of Instructional Resources and Importance placed on Instruction

Example of Survey Items:

- District Curriculum Guide
- Consistency of Math/Science efforts with other school/district information
- Importance school places on Math/Science
- Access to computers/calculators
- Quality of district adopted instructional materials
- Funds for purchasing equipment and supplies
- District/school system for managing instructional resources

2007 Findings:

- Principals reported higher values of "encourages effective instruction" in 2006 than in 2004

  Math $p<0.001$; Science $p<0.0001$
- LoL Participants reported higher values on this scale than non LoL participants

  Math $p<0.05$; Science $p=0.06$

Scale: Effect of District/School Policies and Reward Structures on Effective Instruction

Example of Survey Items:

- Regional curriculum framework
- District/school testing and grading policies and practices
- District/school structures for recognizing and rewarding teacher

2007 Findings:

- Principals reported higher values of "encourages effective instruction" in 2006 than in 2004

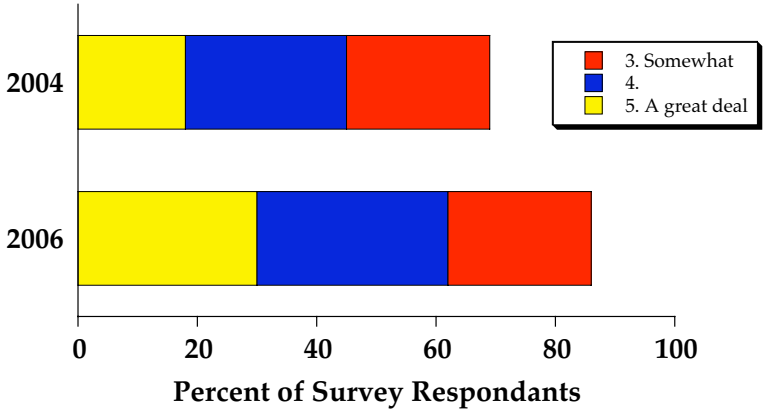
  Math $p=0.09$; Science $p<0.0001$
- No Difference between LoL participants and non participants

Relationship Between MSP Participation and Survey Responses

In keeping with the focus of this report on the effect of participation, the evaluation team examined whether observed changes in principal survey responses could be plausibly attributed to participation in the MSP. As a way to explore this, the evaluation team used statistical models to test whether district-wide participation, as measured by the THASS variable, is related to 2006 survey responses, controlling for 2004 responses. Results of this analysis did not show a statistically significant relationship between participation, as measured by THASS, and the changes in principal survey responses.

As discussed in Chapter 2, the THASS measure only captures one aspect of participation – number of hours of district staff participating in MSP activities. It does not differentiate among the different MSP activities, nor does it distinguish among the type of district staff participating in MSP activities. Although many MSP activities have the potential to contribute to changes in school policies and practices a broad participation measure such as THASS may not capture the impact of participation of principals. Of the

Figure 4-1: Principal-reported Impact of the MSP on School Instructional Practices



MSP activities, *Lenses on Learning* (LoL) is most directly related to principal practice and the one in which principals are most likely to participate. Therefore, the evaluation team examined the relationship between LoL participation and the observed changes in principal survey responses. Analysis compared responses from principals who had completed LoL (36%) with those from principals who had not participated at all (51%) or who had not yet completed the entire LoL series (13%). For this report, the analysis considered only those principals who had completed the original LoL mathematics sessions. This analysis does not consider those principals who had not yet completed the entire mathematics series or the more recently implemented science module. Adjusting for 2004 responses, principals who completed "Lenses on Learning" were more likely to report higher values on some of the same scales as above (see Tables 4-2 and 4-3).

Principals who participated in LoL reported higher values on two of the same leadership skills scales described above—principal knowledge of curriculum and standards and principals' support for teachers to develop professional learning communities. Interestingly, the effect size for LoL participation is greater than the previous analyses that used the THASS measure. For principal knowledge of math curriculum and standards, the effect size controlling for LoL participation is 0.24. The effect size for principal support for teachers to develop a professional learning community is also greater (0.56 for mathematics and 0.41 for science).

LoL participants also had higher values on one of the same scales that showed absolute change—effect and availability of instructional resources and importance placed on math and science instruction. However, in this case, the effect size based on participation in LoL (0.18 for math and 0.19 for science) is not as large as the effect size observed for overall change.

Adjusting for responses in the 2004 survey, in 2006, LoL participants reported 0.6 points higher on a scale of influence of "not at all" to "a great deal" describing the extent that the MSP project overall has influenced mathematics and science instructional practices ($p < 0.001$). In addition, LoL participants scored approximately 0.6 points higher on the scale of individual MSP activities than non LoL participants ($p<0.001$).

Principals also reported on the role of the intermediate units and the MSP staff in supporting their school's improvement efforts in implementing MSP-promoted practices such as data collection and analysis, development of standards, staff development and curriculum. LoL participants were more likely to agree that the IUs and MSP staff support their efforts (effect size of 0.3) than non-LoL participants ($p<0.01$).

Finally, LoL participants showed a higher frequency of reporting that professional development is planned and presented by teachers in the district, as well as that PD being assessed for improvement of teacher practices and student achievement (effect size 0.23, $p<0.05$).

Discussion and Next Steps

The THASS measure of district-wide MSP participation did not prove to be significantly related to changes in principal survey responses. Although a district's overall participation in the MSP might be reflected in the principals' responses, the effect of participation cast broadly might be too diffuse to be detected on the principal survey. A more focused measure of principal participation is therefore worth exploring. Analysis explored the relationship between survey responses and a more direct measure of principal participation, self-reported participation in the LoL series. In more refined future analyses, the evaluation team plans to explore information from the MSP participation database and the MSP staff ratings to obtain a more comprehensive picture of how principals' participation and buy-in relate to principals survey responses.

Using the LoL participation indicator, several scales on the principal survey revealed significant differences between principals who had completed LoL and those who had not participated. These differences were in the desired direction, in that LoL participants' responses are more aligned with the MSP interventions. Moreover, these changes provide evidence for achieving the short- and mid-term outcomes that are expected to precede achievement of the project goals. However, without ruling out other external factors, the differences in survey responses cannot be causally attributed directly to the MSP. For example, it is possible that the principals who are likely to participate in LoL are also principals who are more likely to be open to changing their views and practices and that another factor caused some of the changes observed here. It is also possible that participation in LoL increases the likelihood that principals would endorse survey items consistent with LoL, even in the absence of actual change in knowledge or practice.

As the project continues, the evaluation team plans to continue analyses of the principal survey data. With refined measures of participation, for example, by analyzing data on the actual number of hours of LoL that principals completed. Additional exploratory data analyses may include: using some of the scales developed in the principal survey in achievement models, comparing mathematics- and science-based scales, developing new scales for the items that were added to the survey in 2006, and examining relationships between the principal surveys with the teacher surveys administered in 2004 and 2007.

5. Student Achievement

The ultimate measure of the MSP's impact is student achievement. This chapter describes the various types of student achievement data available for this evaluation, and examines the relationship between measures of participation and student achievement outcomes.

Math and Science Assessments

All public schools in the state administer the Pennsylvania System of School Assessment (PSSA), a standards-based, criterion-referenced assessment used to measure students' attainment of the state's academic standards. Historically, reading and mathematics exams were administered to students in grades 5, 8, and 11. A writing test was added in 2002 for grades 6, 9, and 11, and more recently reading and mathematics tests were added for grade 3. Beginning in 2006, the reading and mathematics PSSA exams were administered to students in grades 3 through 8 and grade 11.

The PSSA uses four performance-level descriptors: the advanced level reflects superior academic performance, proficient reflects satisfactory academic performance, basic reflects marginal academic performance, and below basic reflects inadequate academic performance. These performance levels are reported publicly for all schools in the state, and the evaluation team is collecting this information. In addition, all MSP-participating school districts are requested to provide student-level PSSA results. Some districts have provided this information dating as far back as the 2000-01 school year.

Because PDE does not administer a statewide science exam,[10] the project is annually administering one that is based on the 1995 TIMSS, provided by the PROM/SE[11] Math and Science Partnership at Michigan State University. The MSP has no access to science achievement data from prior to the start of the project, nor from other, non-MSP districts in the state. Moreover, the PROM/SE assessment uses matrix sampling, which does not produce valid student-level scores.[12] Results must be aggregated across groups of students to produce valid and reliable measures.

Student Test Score Database

As mentioned above, the project is building a database of student-level test scores from MSP districts in addition to the statewide school-level PSSA performance level information it is collecting. The database of student scores now contains more than a quarter-million math and science scores. Figure 5-1 depicts the arrangement of these scores by subject area, year, and student cohort. Columns labeled with letters signify cohorts of students. The shaded rows are the years prior to the implementation of the MSP and

---

[10] A PSSA science exam will be initiated in 2008 for grades 4, 8, and 11.

[11] Promoting Rigorous Outcomes in Mathematics and Science Education, http://promse.mspnet.org/.

[12] In matrix sampling assessment designs, students receive different sets of items. This enables broader domain coverage than when all students receive the same items.

MATH	Year	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
2000-2001	Pre-MSP	3441			3943			3849								
2001-2002			5501			5890			6021							
2002-2003				5603			6263			6060						
2003-2004					9061			9686			9182					
2004-2005	1					9039			9973			9042		8202		
2005-2006	2						10326			10858	10611	10206	9593	9425	8900	
2006-2007	3							0			0	0	0	0	0	0
									Gr. 11			Gr. 8	Gr. 7	Gr. 6	Gr. 5	Gr. 4	Gr. 3

SCIENCE	Year	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
2003-2004	Pre					7297			7987			7462				
2004-2005	1						8034			8707			7930			
2005-2006	2							9185			9579			9187		
2006-2007	3								0			0			0	
										Gr. 10			Gr. 7			Gr. 4

Figure 5-1: Math and Science Achievement Scores by Year and Cohort
The numbers indicate the number of student scores available for each year and cohort of students. Columns labeled with letters are cohorts of students. Shaded rows are years prior to implementation of the MSP and unshaded rows are years after implementation. The colored diagonals track the grade levels in which the PSSA is administered. The values in the table indicate the number of student scores currently in the database. Upper table is for math and lower table is for science.

the unshaded rows are years after implementation of the project commenced. The colored diagonals track the grade levels in which the PSSA is administered. The values in the table indicate the number of student scores currently in the database.

Measuring Changes in Achievement Over Time

One way to monitor achievement is to examine a school's performance on a particular PSSA exam. For example, average scores on the Grade 8 PSSA can be examined for trends from pre- to post-MSP. This type of analysis suffers from two serious limitations. First, because the test is given to a different cohort of students each year, it is not possible to rule out the possibility that any observed achievement changes from year to year are due to differences in the student populations taking the exam. Second, PSSA achievement has been increasing statewide over the past few years, and any observed achievement trends in MSP schools might simply be due to the statewide trend.

The following sections present two analytic approaches to address these limitations of the trend analysis: the statewide achievement analysis and student-level achievement models attempt to isolate the estimates of achievement change from statewide trends in test scores; and the latter examines pre-post achievement within cohorts of students in MSP schools.

Statewide Comparison Group

As detailed in the last year's report, the evaluation team created a comparison group of Pennsylvania school districts that matched the MSP districts at the time the project began in 2003 (Williams et al., 2006). Assuming this resulted in well-matched student populations, both groups' achievement trajectories

during the period 2003-08[13] would be expected to be similar if the MSP project had not taken place. However, with the implementation of the MSP project, differences in the student achievement trajectories between the MSP districts and the comparison group districts during 2003-08 can be plausibly attributed to the MSP project. While this analysis is not as rigorous as afforded by an experimental design, where districts are randomly assigned to participate in the MSP or not, it does provide some suggestive evidence of the effect of the project on student achievement. However, there may be unmeasured differences between the MSP and comparison groups, which limits the extent to which this analysis can support strong causal inferences. For the matched comparison group, the evaluation team used Pennsylvania school districts because they utilize the PSSA exam and district-wide results are available publicly. However, because no statewide science exam is administered, this comparative analysis is confined to math achievement.

The comparison group developed through this matching process appears very similar to the MSP districts with respect to many important variables that are often used to predict future achievement. These include: prior PSSA math and reading scaled scores (from 1998 through 2003), the trends of change in those scores, the percentage of students who are in low income families, the percentage of minority students, attendance and graduation rates, the geographic locale (urban, rural, etc.), and the average educational attainment of adults in the community. Overall, this balance is reasonably good; however not all variables balanced well. Notably, while the overall percentage of minority students is well balanced, the percentage of Hispanic students is not. Balance was also not achieved on district size (enrollment) and the total population in the community. Variables that are not well balanced are included as covariates in subsequent linear models.

The PDE discontinued reporting statewide scaled scores in 2004. Thus, after having used a six-year history of math and reading scaled scores in forming the matched comparison group, proficiency levels must be used in models that assess changes in achievement since the project began. This analysis, like the one reported last year, uses the percentage of students scoring advanced or proficient on the PSSA.

Student-Level Achievement Models

For student-level achievement models, the analysis focuses on students for whom the database contains pre-MSP and post-MSP test scores. The best arrangement is to have a score from a student in 2003-04, the last pre-MSP year, and another score three years later. This is exemplified by cohorts G and J in math and cohorts H and K in science in Figure 5-1. This time span is optimal because it provides more time for MSP implementation between the pre- and post-tests, during which the project can have influenced the students' achievement. However, another year of achievement data must be collected before this ideal will be attained. Meanwhile, there are some other cohorts of students that enable pre-post analyses in math across two years of implementation—cohorts F, I, and J. The models discussed later in this chapter use those cohorts. In science there are no cohorts of students with pre-post scores. (But even when pre-

[13] Although the project ends in 2008, results from the 2008 PSSA might arrive too late for analysis for this evaluation. In that case, we will investigate achievement through the year 2007.

post scores become available for some cohorts, student-level analyses will not be possible in science. Analyses will have to be done at an aggregate level, such as the school level, due to the matrix sample assessment design of the science test the MSP is administering.)

Relating Participation to Student Achievement

The statistical models presented here test whether participation, as measured by the THASS variable, is related to achievement after controlling for pre-MSP test scores and other predictors of future achievement.

School-level models comparing MSP schools to matched group of non-MSP schools

Every MSP district has a THASS value. Most non-MSP districts in Pennsylvania had no involvement in the project and are assigned a THASS value of zero. However, some non-MSP districts have participated to a limited extent in MSP activities.[14] As with the MSP districts, their hours are recorded in the participation database. If this participation is affecting student achievement in those districts, it is important to account for this in the statistical models. Therefore, the same method as used for MSP districts was used to compute a THASS value for these non-MSP districts, reflecting the amount of their participation in project activities.

The evaluation team ran statistical models on three grade levels for which scores are available for both 2002-03 and 2005-06 (pre- and post-MSP): grades 5, 8, and 11. Pre-MSP scores and THASS were used as covariates (along with a few variables that did not match well between the MSP and the statewide comparison group). None of these models detected statistically significant relationships between the THASS measure of participation and 2005-06 achievement.

Student-level models examining pre-MSP to post-MSP measures of achievement

For the student-level models, the evaluation team examined three cohorts of students (see Figure 5-1). Cohort F took the 8th grade PSSA in 2002-03 and the 11th grade PSSA in 2005-06; cohort I took the 5th grade PSSA in 2003 and the 8th grade PSSA in 2005-06; and cohort J took the 5th grade PSSA in 2003-04 and the 7th grade PSSA in 2005-06. For each cohort, a hierarchical linear model was used to predict the 2005-06 scores, using the pre-test scores, demographic variables (race/ethnicity, gender, and socioeconomic status), and THASS as covariates, and accounting for the clustering of students within districts. None of these models detected statistically significant relationships between the THASS measure of participation and 2005-06 achievement.

---

[14] Districts not formally designated as MSP participants have always been afforded the opportunity to participate in most MSP-sponsored activities if they cover the cost of their participation.

The evaluation team also explored similar models at the school level (also accounting for school level clustering) using school-level participation measures computed similarly to THASS.[15] These models also did not find any significant effects. Finally, the evaluation team explored using the MSP staff ratings of individual participant buy-in (aggregated to the school and district levels) instead of THASS in the models; those models also showed no significant effects.

Discussion and Next Steps

In summary, using the limited data currently available, the evaluation team found no significant relationships between MSP participation and student math achievement. Over the remainder of the project, additional data will become available that enables recalculating these models for more cohorts of students, in science as well as math, over more years of project implementation, and with a fourth year of participation data. Any refinements that the evaluation team, in consultation with MSP staff, makes to the participation measures will be utilized in future models. The evaluation team will also explore cross-cohort classroom-level models that relate individual teacher participation to achievement. Beyond the various analyses examining the relationship between participation and achievement, the evaluation team will also examine other data sources along the pathways between MSP activities and student outcomes, including survey scales pertaining to practices and policies in MSP schools, derived from the principal and teacher surveys.

[15] Across schools in the MSP districts, the school level variable analogous to THASS ranges from 0 to 64.6, with a median of 9.8.

6. IHE Case Studies

This chapter discusses preliminary findings from the IHE case studies, which were added to the evaluation this year. While much of the information presented in this chapter will be integrated in future data analyses, it is used here to highlight key questions and foreshadow the research directions planned for future analysis.

Contribution of IHEs to MSP Project Goals

The four IHE partners play an important role in achieving the MSP goals. The IHEs are intended to contribute to the project's first long term goal of increasing K-12 students' knowledge of mathematics and science, primarily through the participation of STEM and Education faculty in MSP activities such as the TLAs and CDS, where they expand the content knowledge of in-service teachers. The expected contribution of the IHEs to the project's second goal of increasing the quality of the educator workforce is more direct and occurs through multiple routes. Through participation in MSP activities, IHE faculty members are exposed to the same research-based pedagogy as the K-12 teachers. As a result of this experience, the faculty members may begin incorporating some of these techniques in their own classroom practices. These changes in instructional practice at the IHE level impact undergraduate students and future teachers. Pre-service teachers taking courses from participating STEM and Education faculty are direct beneficiaries of the new techniques and strategies that faculty begin to incorporate in their teaching. In-service teachers who participate in TLAs, which may be facilitated or led by IHE faculty, are also exposed to both content expertise and improved pedagogy. Another expected route to increasing the quality of the educator workforce occurs through CDS, where K-12 teachers are able to participate in in-depth courses on mathematical or science topics which may also increase the quality of the educator workforce. The teacher fellow program may also increase the quality of the educator workforce, by providing K-12 teachers an immersion in the higher education environment. Teacher fellows spend a sabbatical on an IHE campus and participate in coursework as well as assist a faculty member in a course revision. Not only do the teacher fellows benefit, but this fellowship also provides teacher fellows with quality experiences that they can share as a resource with other teachers once they return to their school districts. Finally, IHE participation in all of these activities are intended to contribute to the third long-term project goal of sustainable coordination of partnerships in the IUs that build intentional feedback loops between K-12 and IHE.

In recognition of the importance of the IHEs in the project's long-term goals, the MSP Cabinet allotted additional funding to the evaluation team to expand the evaluation to include IHE case studies. Although data on the IHEs were already being collected as part of the evaluation, the additional resources for case studies have afforded the team an opportunity to examine, in more depth, the following issues:

1) The influence of MSP participation on IHE instructional practices, policies, course content and collaborations among IHE faculty;

2) The influence of IHE participation on MSP partners (K-12 school districts, IUs & other IHE partners) and activities; and

3) The mechanisms within the IHEs that support partnership sustainability.

IHE Case Study Research Questions and Methodology

The case studies are guided by a detailed set of research questions that will provide a framework for examining key partnership issues in the MSP, such as the contribution of IHE partners to the achievement of MSP outcomes and the institutional mechanisms that support that contribution. Because the MSP is a partnership, the evaluation team expects reciprocity among the partners in this interaction and so the MSP's contribution to the IHEs is also examined. A brief summary of the research questions are listed below:

- How has MSP participation influenced the IHEs regarding instructional practices, IHE courses, faculty recognition/ reward policies, and student teacher placement practices?
- How have the IHEs contributed to MSP inputs, activities, outputs and the achievement of specific outcomes (i.e., creation of learning communities, implementation of challenging courses)?
- What are the key IHE institutional factors that enable sustained MSP participation in professional development activities (i.e., content deepening seminars, teacher fellows, and TLAs)?

To address these research questions, the case study methodology includes three main categories of data collection: semi-structured interviews and focus groups; documentation review; and IHE classroom observations. During Year 4, the evaluation team conducted 43 interviews at the IHEs that consisted of interviews with STEM and Education faculty, administrators, and student teacher placement coordinators. Using the IHE case study research questions as a guide, the questions from previous interview protocols were linked to specific research questions and refined as needed to better address the case study research questions. In addition to the interviews already conducted, the evaluation team plans interviews with the project PI, the MSP coordinators, and the K-12 student-teacher placement coordinator, and will conduct focus group interviews with former teacher fellows. Details of the documentation review and the IHE classroom observations will be presented in future reports.

Preliminary Analysis of Case Study Data

In keeping with the major theme of this year's report, focusing on the relationship between participation and outcomes, this preliminary analysis of the IHE interview data focuses on faculty members who have consistently displayed high levels of participation in MSP activities for the previous two years. Using information from the NSF Management Information System (MIS), the evaluation team tracked the numbers of participation hours that faculty members listed and categorized faculty members as either high-participating or medium-to-low-participating. In addition to the MIS data, information from previous IHE interviews was used to make the final determination of high-participating faculty members. For this initial phase of the IHE case studies, interview notes from high-participating faculty members were reviewed and reported. We focus on this subpopulation because these faculty members represent the most active members of faculty at each IHE and are most likely the ones who have consistently participated in MSP activities and may offer important insights into its progress.

This preliminary analysis of IHE interview data examines the variation in responses of high-participating faculty members within and across the IHEs; and the institutional factors that enable high participation among the faculty. Using the IHE case study research questions as the guide, analysis focused on the responses to a subset of interview questions that address the influence of MSP participation on instructional practices, relationships with K-12 districts and other IHE faculty (within and across IHEs) and future impact that MSP participation may have. Questions of partnership, including partnership sustainability and institutional impacts of the MSP, overlapped with many of these topics, and the analysis includes some of these interview questions as well. The sections below highlight some of the findings and themes that emerged from this interview data.

Common Factors Among High-Participating Faculty Members

Determining whether faculty responses differed among IHEs was one of the initial steps in the analysis. In instances where there were differences in faculty responses, the subsequent question was to what extent could these differences be explained by institutional context? Overall, high participating faculty members, across all four IHEs, were similar in their responses to the subset of interview questions that were targeted. The majority of the high-participating faculty shared similar views on MSP activities, their impacts on instructional practices, collaborations among IHE faculty, and student learning. There was some variation among faculty along the issue of partnership, which is discussed in more detail below. Variation along this dimension may not be surprising given the differences in how each IHE became involved in the MSP. As the analysis continues, the evaluation team will explore the relationship between partnership views and motivation for involvement in the MSP in more detail.

Participation in MSP activities

Across the IHEs, all of the high participating faculty members had participated in at least two MSP activities and many had participated in three to four activities. For example, many of them participated in at least one of the TLAs and had either supervised a teacher fellow, developed a CDS, participated on a leadership team, or had been involved in a working group. When asked what they valued most about the MSP activities, many faculty members responded that bringing different people together and getting input from others on issues related to teaching was invaluable. This may not be surprising given that many faculty members also reported having little interaction with each other or having few discussions with their peers about pedagogical issues prior to their participation in the MSP. In essence, while there may have been an existing culture of communication around research at the IHEs, the MSP has created a space for faculty members to begin communicating around issues of teaching and learning.

This increased communication around issues of teaching and learning appears to have moved into practice, as all of the faculty respondents reported changes in their teaching practices. These changes ranged from minor shifts in practice, such as increasing the "pause time" after questioning students, to

incorporating the 5 E model[16] in teaching undergraduate STEM courses. In addition to changes in instructional practices, faculty also discussed the increased awareness of the K-12 culture as an important impact of participation in MSP activities. Faculty reported having a much better appreciation for the pressures of K-12 teachers and their day-to-day struggles. What is less certain from the faculty interviews is the impact on students. Despite expectations that improved teaching practices by IHE faculty will lead to improved learning in undergraduate students, few faculty members could comment on student impact. Some reported that students seemed to have a better sense of the material they are exposed to. However, none could offer more than anecdotal evidence about student impact at this point.

Finally, regarding the question of sustainability of MSP activities, the teacher fellow program was often cited as being sustainable. Reasons for this include the benefits for teachers and faculty as well as the opportunity to develop stronger relationships. Most faculty respondents stated that the CDS had the potential to be sustained, but changes in the way that they are implemented and marketed would need to be undertaken. Many faculty members described the TLAs as among the more valuable MSP activities. Some faculty questioned the extent to which external factors, such as the expense required to support teachers' attendance at the TLAs, and their more formal structure, might affect their sustainability. Finally some faculty felt that the TLAs, in their current configuration, will have run their natural course and, if continued, may evolve to a slightly different structure.

Partnership building

One of the research questions explores how IHE participants have changed in their roles and perceptions of themselves as partners over the course of the MSP. Thus, a key question in the interview protocol is the definition of partnership. Most high-participating faculty members used the term "mutually beneficial" in their description of partnership. However, responses to questions related to partnership building revealed some variation among faculty at the different IHEs. At one end of the spectrum, there were responses from faculty members indicating an overriding sense of isolation and not truly feeling like part of the partnership. At the other end of the spectrum, faculty members reported feeling that partnerships were beginning to grow as a result of committed individuals.

Few faculty members could point to examples of partnerships directly with K-12 districts. However, many faculty members, especially those involved in TLAs and the math/science leadership teams/dinners reported strong relationships with the MSP Coordinators and other MSP staff, who are employees of the IUs and are K-12-based even though they do not work for one specific K-12 district. These interactions were, in many cases, lauded by the faculty members. They also could describe interactions with individuals in the districts, which were suggestive of partnership formation at the district level, but rarely could describe partnerships at an institutional level. For most faculty respondents, teacher fellows provided the best examples of partnerships with K-12 districts. However, for

[16] The 5 E model is a teacher instructional model for introducing new topics to students based on the stages of the learning cycle. This sequence consists of engage, explore, explain, elaborate and evaluate.

many it was not clear how long this relationship would last once the teacher fellows returned to their districts.

Regarding partnerships within IHEs, most faculty respondents could point to increased interactions among faculty members of different disciplines as being important activities for intra-IHE partnership building, and for the most part these were attributable to the MSP. A concern noted by many faculty members is the extent to which these examples of partnerships within IHEs often hinged on a single individual.

Regarding inter-IHE partnerships, many faculty respondents stated that they found value interacting with other individual faculty members, but they were not very optimistic about the sustainability of these interactions. Lack of time and competing interests were often suggested as explanations for why few tangible products were generated from these interactions.

High-participating faculty members' responses were similar when asked about things that have strengthened the partnership with the MSP. Many responded that the individuals involved – individuals who have been committed to participating in activities, willing to consider different approaches to teaching, and willing to work with others to build relationships – were the ones who strengthened the partnership. Faculty responses regarding partnership challenges varied by IHE along a couple of themes. For example, some faculty reported that the challenges to the partnership were largely based on their institutions, specifically its organizational structure and tenure and promotion policies. Other faculty respondents described the partnership challenges as largely due to cultural differences between higher education and K-12. These discussions focused on commonly held preconceived notions by K-12 teachers and IHE faculty and the challenges they create in trying to relate to each other. Finally, faculty members from another IHE were less focused on the cultural differences as barriers to partnership, but did mention factors ranging from logistics to getting upper administrators more involved and supportive of MSP efforts.

Institutional Factors that Facilitate High Participation by Faculty Members

Faculty participation in initiatives beyond their teaching, scholarship, and service responsibilities requires a substantial commitment. The majority of faculty members' decisions to participate in MSP activities may result from their recognition of the benefits to themselves and their students. However, IHEs can also contribute to the likelihood of high participation by providing a favorable environment that supports the interests of the faculty members. High-participating faculty members tend to recognize these institutional factors and verbalize their importance in supporting their choices for how, and how much, to participate in the MSP. The next section discusses some of the institutional factors that enable or facilitate high participation by faculty members. Similar to comments on individual factors that contribute to high participation among faculty members, there were no major differences in the high-participating faculty responses by IHEs to the subset of interview questions focused on institutional factors.

A pre-existing culture of cross-discipline collaboration

One institutional factor that facilitates higher levels of participation in MSP activities is a culture of collaboration. Some high-participating faculty members reported that there was an inclination to work together prior to the MSP. In the sample, evidence of this culture of collaboration included "brown bags" and colloquia geared towards sharing ideas in a constructive rather than critical manner. Less formal forms of the culture of collaboration were also mentioned, such as frequent communication among staff and informal networks of colleagues supporting each other. These types of collaborations were often found among small faculty since "people within the division know each other and help others." However, the benefits of familiarity bred by small staff size may be balanced or even outweighed by the stresses associated with small institutions. Another difficulty resulting from small staff sizes was the lack of professors to cover courses, making course load reduction in reward for MSP activity an unrealistic option for administrators.

An institutional focus on teaching

An institutional focus on high-quality teaching, which is present at all of the IHEs involved, seems to enable high faculty member participation in MSP activities. As one respondent put it, "The key is to be at a place that cares about student learning." Although the reasons for this relationship were not explicitly stated, some can be extrapolated from the interviews.

High-participating faculty members remarked that participating in the TLAs helped them to become better teachers and "grow as professional[s]". In fact, many of the high-participators noted that their teaching had improved since their MSP participation. Some respondents spoke of positive student comments received informally as a result of their improved teaching. One faculty member commented, "Because the discussion is more inquiry-based, students feel they have taken more control of their learning." It can be inferred that these faculty members value the help provided by the MSP in advancing the quality of their teaching. And that their value is, at least in part, because teaching is valued at their institution and counted towards professional advancement. However, respondents also noted that some teachers who participated in the MSP saw their student evaluation scores decline. According to one respondent's interpretation of students' comments, "students are challenged more than they had been in the past, and they [did not] always like that." Although MSP participation may be leading to improved pedagogy, poor student evaluations could negatively impact tenure and promotion decisions and "might be frightening for junior staff that are tenure-track." Therefore, the respondent went on to note, the "MSP challenges us to think more about good teaching and how to measure that."

An institutional emphasis on quality teaching may lead to faculty member participation because, like in K-12 districts, institutions with a teaching focus provide an environment in which the pedagogical approaches espoused by the MSP activities can be more readily implemented. On the other hand, institutions that focus on teaching also tend to have heavy teaching course loads, leading to greater time constraints. High-participating faculty members often noted the value of course reductions in exchange for MSP participation. However, they also recognize the financial implications that make this option unlikely without additional funding to cover additional staff.

An institutional focus on teacher certification and graduate degrees in education

High-participating faculty members noted the compatibility between the goals of the MSP and their IHE's goal to provide a strong nationally-recognized teaching certification program. Many interviewees, including administrators, recognized that participation in the MSP had afforded their institution's teacher certification programs more regional recognition among K-12 districts, and more national recognition among peer institutions. This recognition of the value of MSP involvement facilitated faculty participation in MSP activities. An institutional focus on teacher certification supports the goals of the MSP to improve the qualifications of teachers and affords their faculty members a stronger argument for counting their MSP activities towards the criteria for success within the institution, discussed further below.

Credit towards tenure and promotion

Even within institutions that value teaching, scholarship demonstrated through publishing in peer-reviewed journals remains the gold standard for gaining tenure and promotion. Faculty members repeatedly reported that participation in MSP activities was a risk because of the time it takes away from research. Consequently, faculty respondents said that institutional support of a broad definition of scholarship that includes MSP activities is a facilitating factor for high participation. However, respondents most often mentioned this as a goal rather than a reality.

Respondents across IHEs varied in their reports both of *whether* MSP activities are counted for promotion (i.e., teaching, service, or scholarship) and (*if so) in which categories* the MSP activities are considered. Most faculty members were able to count MSP activities towards their service requirements. These faculty members acknowledged that it was good, at a minimum, to be able to list their MSP participation as service, but that, in reality, service opportunities were not difficult to find. Other respondents reported that MSP activity does not count towards any of their contractual requirements, unless work from it is published in a peer reviewed journal.

Respondents who received credit for MSP participation towards tenure reported institutional tenure and promotion criteria based on the Boyer model, which values a broader definition of scholarship (Boyer, 1990). However, mere participation in MSP is not considered scholarship. For scholarship credit, faculty members were required to apply the knowledge and skills from their experience to professional presentations to peers and colleagues or publications. Many high-participating respondents voiced their goal of working within the existing system for promotion by demonstrating their MSP participation through published works. For some, however, there was uncertainty as to whether or not pedagogical scholarship carried as much weight as more traditional research publications. In many cases there was limited discussion and no written documentation concerning policies for considering MSP activities in tenure and promotion.

Strong support for the MSP from leaders at the departmental and institution level

High participation in the MSP is a challenge without the support of the IHE's administrative leadership because faculty members' primary goal is gaining promotion and tenure. Since promotion decisions are made at the highest administrative level within an institution, the support needs to come from that level as well as the departmental level. Therefore, the evaluation team concludes that support from administrators, as reflected through promotion criteria favorable to MSP participation, is the ultimate facilitator to sustained participation in the MSP. According to many high-participating faculty respondents, this support is not as strong as desired.

Respondents suggested that administrators would respond if made aware of the tangible value brought to their institutions by the MSP, particularly through student satisfaction and national and regional recognition that indirectly increases enrollment and benefits the institution financially. Respondents suggested that data showing a positive impact of the MSP on student satisfaction, faculty teaching, research, and academic programs would help to bolster the evidence that the MSP is contributing institutional improvement. This information could be used in recruitment materials to help draw students as well as to attract faculty members to participate in the MSP.

Next Steps in IHE Case Study Analysis

The findings reported in this chapter are preliminary, based on a first level of analysis of Year 4 IHE case study data. In the coming year, additional interview and observation data will be collected. The evaluation team also will collect additional data from K-12 district staff members who interface with the IHEs and teacher fellows who have collaborated with IHE faculty members. Additionally, the impact of the MSP on teaching will be explored through surveys of curriculum revisions and document reviews of course syllabi. All of these data sources will feed into additional analyses and each of the research questions through comments from high-participating faculty members to better understand the MSP participation. Additionally, a comparison of responses of high-participating and low-to-middle-participating faculty members will allow for analysis of differences in perceptions and practices between these groups. Finally, the evaluation team plans to examine the nature of the changes in IHE policies and practices over the course of the MSP project.

7. Conclusions and Issues for Consideration

This chapter highlights key findings across the Year Four evaluation, especially those related to the focus on participation that is a major theme of this report.[17] Previous chapters presented data and findings from specific instruments and evaluation activity (i.e., K-12 and IHE case studies, principal survey, and student achievement). This chapter draws on findings across instruments and activities to reflect collective thinking about issues for consideration in Year Five, and ultimately, beyond the funded life of the project into a sustainable future. The chapter begins by discussing some challenges to educational reform in large scale efforts, notably time management and curriculum and teaching issues. The remaining sections of the chapter discuss larger themes that are emerging from the work of the SW PA MSP: leadership and advocacy, adaptation and sustainability, the achievement horizon, and the nature of educational reform.

Supporting Change in a Large Scale Project

The SW PA MSP is a large-scale, comprehensive effort. The project involves 53 school districts representing 39% of the total number of public school districts in the 11-county region (an extended population center emanating from Pittsburgh). Those districts are diverse in their financial, educational, and geographic contexts, covering urban, suburban, and rural districts of varying size and wealth. While it is important to focus evaluation efforts on examining mid- and long-range outcomes (see the logic model presented in Chapter 2), it is also necessary to credit the project with its ambitious scope and the documented implementation of the intervention strategies. Few, if any, reform initiatives in the region, to date, have had the extensive and deep reach of the MSP. The scope of the project will hopefully ensure a continuing infrastructure for change and growth just as its predecessor organization, the Math Science Collaborative, had developed a strong foundation to support MSP development and implementation.

Educational reform is a process that presupposes that individuals, and through collective effort, institutions, become aware of the need to change, are equipped with the knowledge to change in an informed way, commit to change, and commit to evaluate and further calibrate that change. Each phase of that cycle is replete with challenges that hinder progress. Major initiatives such as the MSP are designed to address each element of the cycle and successfully meet these large challenges. Educational reform literature is steeped with examples of good intentions and well-designed and implemented change efforts that failed as they attempted to move to scale (Bowman, 2005; Center for Mental Health in Schools at UCLA, 2005; Christina & Nicholson-Goodman, 2005; Taylor, Nelson & Adelman, 1999).

This MSP has faced many of the challenges identified in the literature. Building an environment supportive of reform at both the K-12 or IHE levels is necessary for reform to take hold and grow. A number of factors are apparent in the data across the spectrum of K-16.

---

[17] See Chapter 2 for a thorough discussion of the use of an adjusted participation variable (THASS) as the basis for K-12 participation measure.

Time Management

Time management issues seem most apparent at the more structured and time-sensitive K-12 level. However, as discussed in Chapter 6, it is also an issue in IHE, though manifested somewhat differently. At the K-12 level, the case study and principal survey data indicate an ongoing struggle for districts to be able to meet the professional development requirements of the MSP. Schools struggle to enact the on-site professional development, in part due to reductions in the number of hours available for training, and other disruptions to the planned professional development schedules. Planned sessions are often supplanted by competing priorities and professional development needs.

IHE partners struggle with time management in a somewhat different way. While individual faculty experience more independence in deciding how much time to engage in MSP activity than do their K-12 counterparts, there is less institutional support and acknowledgement of MSP involvement as bona fide activity. As discussed in Chapter 6 and prior reports (Williams, et al., 2005, 2006), the incentive and reward structures in higher education tend to view faculty work with the MSP as "service," rather than as scholarship or teaching. Scholarship and teaching are more highly valued in the promotion and tenure process. While this issue has been more directly addressed over the last two years with evidence of written policy statements by Deans at the IHE campuses, faculty still express serious doubt about how these policies will be put into practice in existing promotion and tenure decision-making.

Curriculum and Teaching

Case study data from both K-12 districts and the IHEs indicate that some changes in curriculum and teaching may be attributable to the MSP. Teacher leaders and some teachers who have received training at on-site academies have demonstrated changes in the way they approach what they teach as well as the way they implement lessons. Further, principals, especially those who have attended the *Lenses on Learning* training sessions offered through the MSP report a deeper understanding of what constitutes good mathematics and science lessons and a greater confidence and ability to support it. Administrative supervision and support are important conditions to sustain teacher changes in curriculum development and instruction.

So too, a number of IHE faculty report changes in how they plan and deliver their courses. Although these changes may be in line with the overall MSP goals, there is a concern that faculty members will not be recognized for these efforts because they have not resulted in publications. Publications, rather than good pedagogy, are the most highly valued currency for promotion considerations, presenting challenges to all faculty members but especially those in the STEM disciplines where research expertise is paramount. The value individual faculty members place on improved teaching is not reflected in the institution or the academy as a whole as evidenced in policies and practices related to promotion.

Leadership and Advocacy

Leaders, as discussed earlier, can provide important contextual supports to establish and sustain reform or can allow it to wither from a lack of support. As discussed in Chapter 4, principals who have

completed *Lenses on Learning* training report higher values on a number of scales related to mathematics curriculum and standards, the importance placed on mathematics and science instruction and the effect of quality resources, and the importance of supporting continuous professional development for their teachers. So too, these principals report a commitment to providing a more supportive environment for reform-based pedagogy. However, findings indicate a wide variation of MSP implementation across districts, and even some of those with LoL-participating principals are low on participation measures. Just as teacher leaders must address how best to implement mathematics and science reform in K-12 schools in the midst of other priorities and reform measures, so too, principals must struggle with similar issues in their support of educational reform. Clearly, there has been consistent evidence across the first four years of MSP implementation that some level of advocacy is required in schools to champion a reform initiative such as the MSP. At times this advocacy has started first with teachers, at other times, with administrators. However, there is growing evidence to suggest that administrative support, and perhaps, advocacy, is required for such an intensive effort to be sustainable.

At the IHE, similar issues emerge. The evaluation team has reported in the past that partnership, especially at the IHE level, is forged first and foremost, via individual faculty. While an institution may indicate a policy of partnership in the MSP, this is often translated through a few faculty members who advocate for the efforts. Across the four years of the MSP to date, there has been considerable effort by the Deans of participating IHEs to forge relationships with one another and articulate their support of the MSP and its sustainability. One example of this has been the Dean's Dinners and a joint policy statement regarding the importance of faculty involvement, especially in relation to promotion and tenure. Even with these efforts, plans for sustainability among the IHE partners seems less tangible among faculty than among some K-12 school and IU partners, and issues of leadership and advocacy are yet to be resolved.

Adaptation and Sustainability

The annual MSP implementation plans always included an awareness of the need to promote sustainability among partners beyond the NSF funding period. This is evidenced in a number of ways from the planned expansion beyond original partners to the insistence by the Cabinet and PI that K-12 schools, IUs, and the IHE partners commit to an escalating schedule to assume payment for infrastructure support across the life of the grant. One aspect of the sustainability plan, according to the PI, is a similarly escalating local control of the reform initiative through adaptation. The implementation plan and requirements were communicated and monitored more rigidly during the early years of the project when partnering sites were beginning implementation internally. Implementation in later years (especially in Years Four and Five) is designed to support more experimentation and accommodation of institutional adaptation, which in turn is intended to facilitate integration of reform within organizations.

Over time, K-12 and IHE partner institutions change, as do priorities and resources. K-12 changes sometimes occur rapidly, especially during an era of administrative turnover. As reported earlier in this document, the principal survey population changed significantly from 2004 to 2006 likely as the result of turnover. This is further exacerbated by changes in teacher assignments across years, diluting the training effect within a particular grade level, discipline, or school. Districts' adaptation of the implementation

plan helps to address these issues. However, the project and the schools struggle to balance acceptable versus fatal adaptation for meeting the goals of the project at particular sites.

Higher education faculty work more independently than do K-12 teachers. Consequently, adaptation is less problematic for individual faculty at the IHE level. Adaptations to curricula or pedagogy are much more likely to stay with that particular individual and there are fewer opportunities for dissemination across colleagues. Moreover, higher education faculty do not as often work from a common syllabus and incoming faculty may choose to teach courses quite differently than their predecessors.

Cultural issues, both in K-12 and IHE, also present challenges in adaptation and sustainability across the partnership. The K-12 institutional environment is designed to be amenable to public policy change through state and local control. There has been visible evidence of how much both federal legislation and state mandates impact how education is organized and delivered as a result of various state standards and assessments and other policies stemming from No Child Left Behind. The more recent press in educational accountability measures have created an environment where change is expected. This is less evident in IHEs currently, though new efforts are afoot to bring similar measures of accountability to bear in the post-secondary educational arena (Fisher & Hebel, 2007; AASCU, 2007). In some ways, the higher education culture is specifically designed to resist organized and systemic change. The tenure system is, in part, designed to assure an unregulated organization and dissemination of knowledge as determined by faculty "entitled to freedom in the classroom in discussing their subject" (AAUP, 2006).

The Achievement Horizon

As similarly reported in the Year Three report (Williams, et al., 2006), changes in student achievement have not been observed in statistical models of the data available thus far. Benchmark and other data (PSSA group-level data, Profile of Mathematics and Science Indicators data, number of courses revised by IHE, students passing IHE courses, etc) *do* show encouraging changes in short-term and mid-term outcomes that are generally accepted as necessary precursors to changes in achievement, although in many cases these changes cannot be causally attributed to the MSP. The lack of change in PSSA mathematics scores at the individual student level raises a number of issues or questions: For example, the PSSA exam may not be sufficiently sensitive to the changes induced by the MSP, or it may require more time for the changes to affect PSSA results broadly across a significantly large group of students. This is particularly important in schools where only subsets of teachers participate in the MSP.

The evaluation team is in the midst of refining the statistical models being used to examine changes in student achievement. While the THASS participation measure was developed for the purposes of organizing this report, it is not a perfect measure as discussed in Chapter 2. The depth and breadth of data, both in terms of student achievement and in educator participation provides opportunities to explore relationships of specific groups of students and teachers. For example, in a subset of districts that have provided detailed scheduling information for students, there is a plan to explore analyses that focus more directly on students in classes taught by MSP-participating teachers. The evaluation team is engaged in discussions with the PI, MSP staff, and Cabinet members to help to inform and prioritize appropriate avenues for further analysis of achievement data.

An important discussion, already begun by the evaluation team, and most likely to extend to the MSP Cabinet in the next six months, revolves around the complexity and challenge of educational reform and impact on student achievement. The project is most of the way through both an intensive and extensive reform initiative. If changes in student achievement are not yet evident, then important questions must be asked about the feasibility of this type of reform and the measures used to track its impact. The efforts of the SW PA MSP have been research-based and well-designed, administered with care, precision and accuracy, implemented under realistic though not ideal conditions, and monitored using assessments that are similar to those being used across the states. While evidence is strong that change is occurring among K-12 educators and IHE faculty, it is not clear how long it may take to have a measurable impact on student achievement. Many questions regarding the achievement horizon still remain, in particular whether factors beyond the reach of a reform such as the MSP are pushing the horizon further away despite best efforts.

The Nature of Educational Reform

As the findings discussed in this report are considered, especially those about the lack of documented changes in student achievement to date, a number of issues related to the nature of educational reform emerge from this evaluation activity. A renewed focus and accelerated activity for educational reform can be traced to the 1983 report, *A Nation at Risk* (Gardner, et al., 1983). Through the 1980's and 90's, many initiatives were funded to address the concerns outlined in this report, some comprehensive, such as America's Choice Schools and Success for All, and some more focused on a discipline, such as QUASAR mathematics. The renewal of the Elementary and Secondary Education Act in 2001 as *No Child Left Behind* continued the growing trend toward quantitative accountability and educational reform. The last decade of standards-based educational reform, coupled with accountability-driven mandates has generated a burgeoning list of activity in public school systems.

In part, the establishment of the founding organization for the SW PA MSP, the Math Science Collaborative, was motivated by the many mathematics and science education reform efforts underway in the region that were disconnected from each other. This disjointed reform landscape did not support educational change across a region concerned with economic development and revitalization. Although efforts in math and science across the region are more coherent today as a result of the Math Science Collaborative and the MSP, the overall educational reform landscape is even more complicated and disjointed today as a result of the many reform initiatives across levels and disciplines.

The K-12 case studies reveal that many schools are inundated with reform activity. Even among higher-performing schools, there is evidence of reform fatigue with educators feeling pressured from an increased level of public scrutiny and the rollout of initiative after initiative, year after year. Many educators who have lived through decades of reform that have come and gone have adopted a "this too shall pass" attitude that in part manifests in an expectation that reform efforts will, at best, last about three years, and simply fade away as funding or leadership wanes. Even more damaging is the fact that schools facing the greatest educational challenge are often a hotbed for reform activity that is particularly incoherent. Activities and expectations compete with one another, sometimes actually colliding in their requirements. While there is evidence of much activity, the agenda can often be unfocused and

unproductive, further confirming the suspicion among seasoned educators that reform efforts make little or no difference. Repeated episodes of effort without substantive improvement may lead to a sense of relentless failure, making it nearly impossible to mount a sustainable change effort. These issues may, in fact, account for a disturbing finding from the case studies. The lowest-participating districts have included a number of districts that fit the descriptions offered above. They are low-performing, financially and educationally challenged schools that are in many ways laboratories for reform. Unfortunately, there is evidence in each of these cases of this brand of reform fatigue coupled with a sense of expected failure among teachers and administrators.

Of equal concern is the notion that there is a "silver bullet" or "best" practice that will resolve the achievement needs of students. While educators are often quick to caution others to not fall prey to the lure of simple solutions to complex issues, they seem equally compelled to seek change that is less intrusive to their existing practice. As Fullan (2002) and others (Elmore, 2004; Fuhrman, 2003) have suggested, change is never easy when it requires educators to rethink and redesign practice and question deeply held assumptions about students and ourselves. Perhaps the natural proclivity to seek external resolution to and by others also contributes to the expectation that educational reform has a limited life and scope of impact.

These challenging issues are not new nor are they easily addressed. They are not indigenous only to math and science education reform nor are they issues for the SW PA MSP alone. They are ubiquitous and recurring issues that challenge educational reform broadly and remain active hurdles for the MSP.

References

American Association of University Professors (2006) 1940 Statement of Principles on Academic Freedom and Tenure with 1970 Interpretive Comments. [available at http://www.aaup.org/AAUP/pubsres/policydocs/1940statement.htm].

American Association of State Colleges and Universities. (2007). 2007 Public Policy Agenda. Washington, DC: American Association of State Colleges and Universities.

Borman, G. D. (2005) National efforts to bring reform to scale in high-poverty schools: Outcomes and implications. *Review of Research in Education, v29.* Washington, DC: American Educational Research Association, p1-28.

Boyer, E. L. (1990). Enlarging the perspective. In *Scholarship Reconsidered: Priorities for the Professorate*. (pp. 15-16). Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Center for Mental Health in Schools at UCLA. (2005). Systemic change for school improvement: Designing, implementing, and sustaining prototypes and going to scale. Los Angeles: UCLA Center for Mental Health in Schools.

Christina, R. & Nicholson-Goodman, J.V. (2005). Going to scale with high-quality early education. Santa Monica, CA: RAND Corporation.

Elmore, R.F. (2004) School Reform from the Inside Out: Policy, Practice, and Performance. Boston: Harvard Educational Pub Group.

Fisher, K. & Hebel. S. (2007) Key higher–education issues in the states. *Chronicle of Higher Education, v53n18*. Washington, DC: Chronicle of Higher Education, p. A9

Fullan, M. The principal as a leader in a culture of change, Educational Leadership, May, 2002

Fuhrman, S. (2003). "Riding waves, trading horses. The twenty-year effort to reform education." In D.T. Gordon (Ed.), *A nation reformed? American education 20 years after A Nation At Risk* (pp. 7-22). Cambridge, MA: Harvard Education Press.

Gardner, David P., chair. A Nation At Risk: The Imperative for Educational Reform. The National Commission on Excellence in Education, U.S. Department of Education, Washington, DC, April, 1983. ED 226 006.

Le, V.-N., Stecher, B. M., Lockwood, J. R., Hamilton, L. S., Robyn, A., Williams, V. L., et al. (2006). Improving Mathematics and Science Education: A Longitudinal Investigation of the Relationship Between Reform-Oriented Instruction and Student Achievement (Monograph MG-480-NSF). Santa Monica, CA: RAND Corporation.

Lewis, C., R. Perry, and J. Hurd (2004), A Deeper Look at Lesson Study, Educational Leadership, February 2004, p.18-23.

Taylor, L., Nelson, P., and Adelman, H. S. (1999) Scaling up reforms across a school district. *Reading & Writing Quarterly 303(23).*

U.S. Department of Education. (2006). Common Core of Data: Information on Public Schools and School Districts in the United States. Retrieved March, 2007, from http://nces.ed.gov/ccd/

Vernez, G., Karam, R., Mariano, L., T., & DeMartini, C. (2006). Evaluating Comprehensive School Reform Models at Scale: Focus on Implementation (Monograph MG-546-EDU). Santa Monica, CA: RAND Corporation.

Williams, V. L., Pane, J. F., Tananis, C. A., & Olmsted, S. S. (2005). Math Science Partnership of Southwest Pennsylvania: Year Two Evaluation Report (Working Paper WR-270-EDU). Santa Monica, CA: RAND.

Williams, V. L., Pane, J. F., Tananis, C. A., Olmsted, S. S., & Ciminillo, C. (2006). Math Science Partnership of Southwest Pennsylvania: Year Three Evaluation Report (Working Paper WR-396-EDU). Santa Monica, CA: RAND.