



**Mathematics and
Science Partnerships:
Summary of
Performance Period
2012 Annual Reports**

**Analytic and
Technical Support
for Mathematics
and Science
Partnerships**

Contract # ED-CFO-10-A-0024

December 1, 2014

Prepared for
Patricia O'Connell Johnson
Irene Harwarth
U.S. Department of Education,
400 Maryland Ave, SW
Washington, DC 20208

Prepared by
Abt Associates Inc.
Ellen Bobronnikov
Rebecca Gotlieb
Peter Honnef
Katherine Murphy
Sarah D. Sahni
Alex Silverman

150 Wheeler Street
Cambridge, MA 0213

Contents

Executive Summary	1
Performance Period 2012 Mathematics and Science Partnerships	1
Professional Development Content, Models, and Activities	3
MSP Evaluation Designs and Outcomes.....	4
Conclusions.....	5
Chapter 1: Introduction.....	7
The Mathematics and Science Partnership Program.....	7
Report Overview and Analytic Approach.....	9
Report Organization.....	11
Chapter 2: Characteristics of MSP Projects and Participants.....	12
Sources and Amounts of Funding.....	12
Organization and Partnerships	13
!Unexpected End of Formula	
Methods of Selecting Participants.....	16
School Levels and Types of Participants Served	16
Chapter 3: Professional Development Content and Activities	19
Professional Development Content of MSP Projects.....	19
Professional Development Models	21
Primary Academic Year Professional Development Activities	23
Chapter 4: MSP Evaluation Designs and Outcomes.....	26
Evaluation Designs	26
Measures Used in Evaluations	27
Evaluation Findings	30
Chapter 5: Highlights from MSP Projects with Rigorous Designs.....	33
Chapter 6: Summary and Conclusions.....	64
References	66
Appendix A: Review of Projects with Rigorous Designs	67
Methodology Used for Review	67
Attrition.....	69
Baseline Equivalence	70
Quality of Measurement Instruments.....	71
Relevant Statistics Reported	72
Summary	73
Appendix B: Criteria for Classifying Designs of MSP Evaluations.....	75
Appendix C: 2012 State MSP Appropriations.....	78

Executive Summary

Improving students' achievement in mathematics and science will be critical to maintaining the nation's competitiveness. Research on teacher quality has demonstrated that one of the strongest indicators of students' academic success is the competence and capability of their teachers (Clotfelder, Ladd, & Vigdor, 2007; Hanushek & Rivkin, 2010; Hill, Rowan, & Ball, 2005; Rivkin, Hanushek, & Kain, 2005). Thus, education improvement efforts around the country are increasingly focused on the teacher as the most powerful agent of change for improving student learning.

As the limitations of short-term professional development opportunities for teachers have been recognized, there has been widespread interest in sustained university partnerships with local school districts to offer rich professional learning opportunities for teachers and administrators. The U.S. Department of Education's Mathematics and Science Partnership (MSP) program funds collaborative partnerships between high-need school districts¹ and mathematics, science, and engineering departments at institutions of higher education (IHEs) for the purpose of providing intensive content-rich professional development to teachers and other educators, thus improving classroom instruction and ultimately student achievement in mathematics and science.

Implemented under the *No Child Left Behind Act of 2001, Title II, Part B*, MSP is a formula grant program to the states, with the size of individual state awards based on student population and poverty rates. The states then award the funding on a competitive basis to local partnerships. Federal support for MSP increased substantially from the program's inception in FY 2002—from \$12.5 million to \$100 million in FY 2003, when MSP became a state-administered formula grant program. Funding has since increased further, ranging from \$150 to \$182 million awarded to local partnerships each year. In FY 2012, grants to states totaled \$150 million.

Performance Period 2012 Mathematics and Science Partnerships

This report presents an overview of the MSP program during Performance Period 2012 (PP12),² including the characteristics of MSP projects and participants; the professional development content, models, and activities of the projects; and the MSP projects' evaluation designs and outcomes.

Amount of Funds

In PP12, federal MSP resources totaling \$150 million were distributed to the 50 states, the District of Columbia, Puerto Rico, and U.S. Island areas.^{3,4} State grants ranged from approximately \$745,000 up to \$18 million with an average of \$2.8 million and a median of \$1.7 million. In turn, the states funded

¹ The term "high-need" is not explicitly defined in the statute for the Mathematics and Science Partnership program. Each state educational agency is responsible for conducting a needs assessment to determine the highest priority for these professional development funds and for defining high-need for its grant competition.

² Performance Period 2012 (PP12) refers to the period between October 1, 2012 and September 30, 2013. Information in this report is based on the Annual Performance Reports submitted regarding this performance period.

³ The American Virgin Islands, Guam, the Northern Mariana Islands, and Samoa pool their MSP funds as part of their consolidated budget. They are not required to submit annual performance reports to the MSP program, so their activities are not reflected in this report.

⁴ One half of one percent of total funds is reserved for program administration.

a total of 488 local MSP projects, with local grants ranging from approximately \$20,000 to over \$5.8 million with a median project grant of approximately \$213,000 and a mean of approximately \$345,000. As shown in Exhibit ES.1, most projects (83 percent) received \$500,000 or less in funding. In addition to these federal funds, some local projects reported receiving supplemental funding from other federal and non-federal sources.

Exhibit ES.1: Sub-Grant Budgets from State MSP Grants, Performance Period 2012

Project Budgets	Percent of Projects (N=487)
\$100,000 or less	16%
\$100,001 to \$200,000	32
\$200,001 to \$500,000	35
\$500,001 to \$1,000,000	13
\$1,000,001 or more	3

Source: Annual Performance Report item I.A.6
The non-response rate was <1 percent in PP12.

Participant Selection

When asked about the main goal of their MSP project, nearly three-quarters of projects (74 percent) indicated that it was to improve individual teachers’ content knowledge, while only 2 percent reported that it was training teacher leaders who would in turn train other teachers. Twenty-two percent of projects reported that both goals were equally important, indicating that most projects that train teacher leaders also train individual teachers.

Characteristics of Project Participants

Over three thousand faculty members from IHEs were involved with MSP projects in PP12, with an average of 7 IHE faculty members per project. Projects are required to establish direct interactions between K–12 teachers and IHE faculty members in mathematics, the sciences, engineering, or technology. Additionally, nearly two-thirds of the projects (64 percent) reported working with faculty members from education departments within IHEs.

Over 40,000 elementary, middle, and high school teachers, coaches, paraprofessionals and administrators participated in MSP projects in PP12. The number of these participants served by individual MSP projects ranged widely from 6 to 1,382, with typical projects serving 47 participants. These participants, in turn, taught over 2.3 million students.

Eighty-six percent of MSP participants were regular classroom teachers of core mathematics and/or science content. In order of prevalence, the remaining 14 percent of participants included special education teachers, school administrators, gifted and talented teachers, ELL teachers, math coaches, paraprofessionals, and science coaches.

School Levels

MSP projects are free to select the grades or school levels in which they provide professional development. Among the individuals participating in MSP activities, 47 percent were employed at the elementary school level, 28 percent at the middle school level, and the remaining 25 percent at the high school level. Nearly three-fourths of projects (74 percent) targeted multiple school levels (i.e.,

some combination of elementary, middle, and/or high school); 38 percent served participants from all three school levels.

Professional Development Content, Models, and Activities

Professional Development Content and Processes

In PP12, 36 percent of MSP projects provided professional development in both mathematics and science; 40 percent provided professional development in mathematics only; and 24 percent provided professional development in science only.

Most MSP projects addressed multiple content areas and processes, both within and across disciplines. Across elementary, middle, and high school, scientific inquiry was a frequently cited process taught (92 to 98 percent of projects that addressed science), while physical science was the most frequently cited science content area (69 to 73 percent), and chemistry was the least frequently addressed content area (41 to 46 percent). Technology was addressed by approximately two-thirds of projects (64 to 71 percent). In mathematics, problem solving was a frequently addressed process taught across levels (83 to 88 percent of projects that addressed mathematics); number and operations was the most commonly addressed content area in elementary school (80 percent); and algebra was the most frequently addressed content areas in middle and high school (74 and 73 percent, respectively). Calculus was the least frequently addressed topic (4 to 22 percent of projects that addressed mathematics).

Professional Development Models

The MSP program legislation defines a summer institute as a model of professional development that provides intensive learning experiences over a minimum of a two-week period. As shown in Exhibit ES.2, half of projects (50 percent) conducted summer institutes with academic-year follow-up activities.⁵ These projects reported offering a median of 97 hours of professional development. Just 3 percent of projects provided summer institutes only, with no follow-up. The remaining projects (47 percent) offered a broad range of professional development models, including onsite professional development, graduate courses, on-line coursework, and professional learning communities. Most of these projects also included a shorter summer component. The median length of professional development for these projects was 68 hours.

Exhibit ES.2: Median Professional Development Hours, by Professional Development Model Type, Performance Period 2012

Professional Development Model	Percent of Projects (N=481)	Total Median Hours
Summer institute only	3%	80
Summer institute with follow-up	50	97
Various other models	47	68

⁵ Summer institutes are defined in the MSP legislation as providing intensive learning experiences for a minimum of two weeks during the summer. Projects that included summer workshops that were less than two weeks were classified as projects with a focus on academic-year activities.

Source: Annual Performance Report item V.A.1, V.B
 The non-response rate for the model type was 0 percent.
 The non-response rate for the total hours of professional development was 4 percent.

Professional Development Activities

The professional development activities offered by MSP projects focus on increasing teachers’ content knowledge in mathematics and/or the sciences and on enhancing their pedagogical skills. The most commonly reported model for delivering academic-year activities was on-site professional development (69 percent of projects), followed by study groups (14 percent), content coursework at colleges or universities (6 percent), and on-line coursework/distance learning networks (3 percent).

MSP Evaluation Designs and Outcomes

Evaluation Designs

MSP projects reported the primary designs they used to assess program outcomes. Fewer than one percent reported using an experimental design in which teachers, classrooms, or schools were randomly assigned to a treatment or control group. Nearly half of the projects (48 percent) reported using a quasi-experimental design with a matched or non-matched comparison group. The remaining projects reported using less rigorous evaluation designs, such as: single-group design with pre- and post-tests (36 percent); qualitative or descriptive methods only (12 percent); or mixed quantitative and qualitative methods (3 percent).

The MSP program has been advising its grantees on implementing impact evaluations by providing them with the MSP criteria for carrying out rigorous impact evaluations and providing guidance regarding how to interpret these criteria and report on their evaluations. A review of final-year projects was performed to determine the extent to which projects successfully conducted rigorous evaluations to yield findings that could be considered reliable and valid. As Exhibit ES.3 shows, the number of projects with at least one evaluation meeting all criteria increased four-fold from PP07 to PP09 and the proportion of passing projects has continued to rise.

Exhibit ES.3: Final Year Projects that Conducted Rigorous Evaluations and Met MSP Criteria for Rigor, Performance Periods 2007–2012

Projects	PP07	PP08	PP09	PP10	PP11	PP12
Implemented comparison group designs	37	49	65	59	59	71
Included at least one evaluation that met all criteria	4	3	16	15	17	21
Percent of projects with at least one evaluation that met all criteria	11%	6%	25%	25%	29%	30%

Teacher Content Knowledge Outcomes

Federal regulations require that all teachers who receive MSP funded professional development are pre- and post-tested at least once during the life of a project. As shown in Exhibit ES.4, over half of teachers who received professional development in mathematics and science were tested using pre- and post-assessments in PP12 (56 percent in mathematics and 62 percent in science). Sixty-three percent of teachers who were assessed in mathematics and 67 percent of teachers who were assessed in science showed statistically significant gains in their content knowledge.

The most frequently reported assessments of teacher content knowledge in mathematics were nationally normed/standardized tests (67 percent of projects). Projects that did not use nationally normed or standardized content assessments often developed their own assessments for their MSP projects. Approximately one-third of projects (34 percent) used locally developed tests to assess teacher gains in mathematics content knowledge. In science, the most frequently used instruments were locally developed tests (53 percent of projects), followed by standardized instruments (48 percent).

Exhibit ES.4: Percent of Teachers with Significant Gains in Content Knowledge, Among Teachers with Pre-Post Content Assessments, Performance Period 2012

Content Area	Total Number of Teachers Served	Percent of Teachers with Content Assessments	Percent of Assessed Teachers with Significant Gains
Mathematics	23,546	56%	63%
Science	16,167	62	67

Source: Annual Performance Report items VIII.A. 1, 2, 3, 4, 5, 6
Individual teachers who received professional development in both mathematics and science may be included in the number of both science and math teachers.

Student Achievement Outcomes

As shown in Exhibit ES.5, among the 50 percent of students with assessment data in mathematics, 55 percent scored at the proficient level or above, which represents a substantial decrease from the previous year (when 64 percent of students scored at proficient or above). However, among the 29 percent of students with assessment data in science, 69 percent scored at the proficient level or above. This represents a substantial increase from earlier years in the proportion of students with assessment data scoring at the proficient level or above in science.

Exhibit ES.5: Percent of Students Scoring at Proficient Level or Above, Among Students Taught by MSP Teachers and Assessed In Each Content Area, Performance Period 2012

Content Area	Total Number of Students Taught by MSP Teachers	Percent of Students with Assessment Data	Percent of Assessed Students at Proficient Level or Above
Mathematics	1,038,381	50%	55%
Science	787,007	29	69

Source: Annual Performance Report items VIII.B. 1, 2, 3, 4, 5, 6, 7, 8
Students who are taught by teachers receiving professional development in math and science may be double counted.

Conclusions

The MSP program is successfully implementing the requirements of the law. Partnerships are being formed between STEM and education departments at IHEs and high-need local educational agencies, and many of these partnerships also include public or private schools, businesses, and non-profit or for-profit organizations. Teachers are receiving intensive and sustained content-rich professional

development—from college and university faculty partners and other professionals—that integrates mathematics and science content with effective pedagogical strategies. Many of these teachers have the additional advantage of receiving ongoing mentoring and coaching from faculty and master teachers as they begin to implement their new knowledge and practices in their classrooms. Furthermore, many projects are collecting data on what teachers are learning and are conducting rigorous impact evaluations.

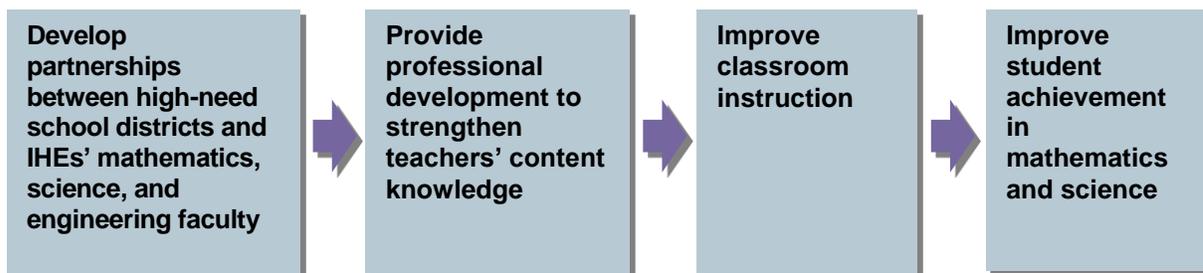
In PP12, over 7,000 local educational agencies (LEAs), organizations, and institutions—involving over 3,000 IHE faculty members—partnered to form 488 projects across the country. Projects served over 40,000 educators nationwide, with each educator receiving an average of 97 hours of professional development, thus enhancing the quality of classroom instruction for over 2.3 million students. Based on this professional development, 63 percent of teachers who were assessed in mathematics and 67 percent of teachers who were assessed in science showed statistically significant gains in their content knowledge. Over half of students taught by MSP teachers scored at the proficient level or above in state assessments in math and science (55 percent and 69 percent, respectively).

Chapter 1: Introduction

Improving students’ achievement in mathematics and science will be critical to maintaining the nation’s competitiveness. Research on teacher quality has demonstrated that one of the strongest indicators of students’ academic success is the competence and capability of their teachers (Clotfelder, Ladd, & Vigdor, 2007; Hanushek & Rivkin, 2010; Hill, Rowan, & Ball, 2005; Rivkin, Hanushek, & Kain, 2005). Thus, education improvement efforts around the country are increasingly focused on supporting teachers as a powerful approach to improve student learning.

The limits of short-term professional development offerings for teachers have been documented, leading to a push for more sustained and focused professional learning for teachers. In efforts around the country to improve mathematics and science learning, there has been interest in supporting partnerships between university faculty and local school districts in order to offer rich professional learning opportunities for teachers and administrators. The U.S. Department of Education’s Mathematics and Science Partnership (MSP) program funds collaborative partnerships between high-need school districts and mathematics, science, and engineering departments at institutions of higher education (IHEs) for the purpose of providing intensive content-rich professional development to teachers and thus improving classroom instruction and ultimately student achievement in mathematics and science (see Exhibit 1).

Exhibit 1: Conceptual Model of Mathematics and Science Partnerships Program



The Mathematics and Science Partnership Program

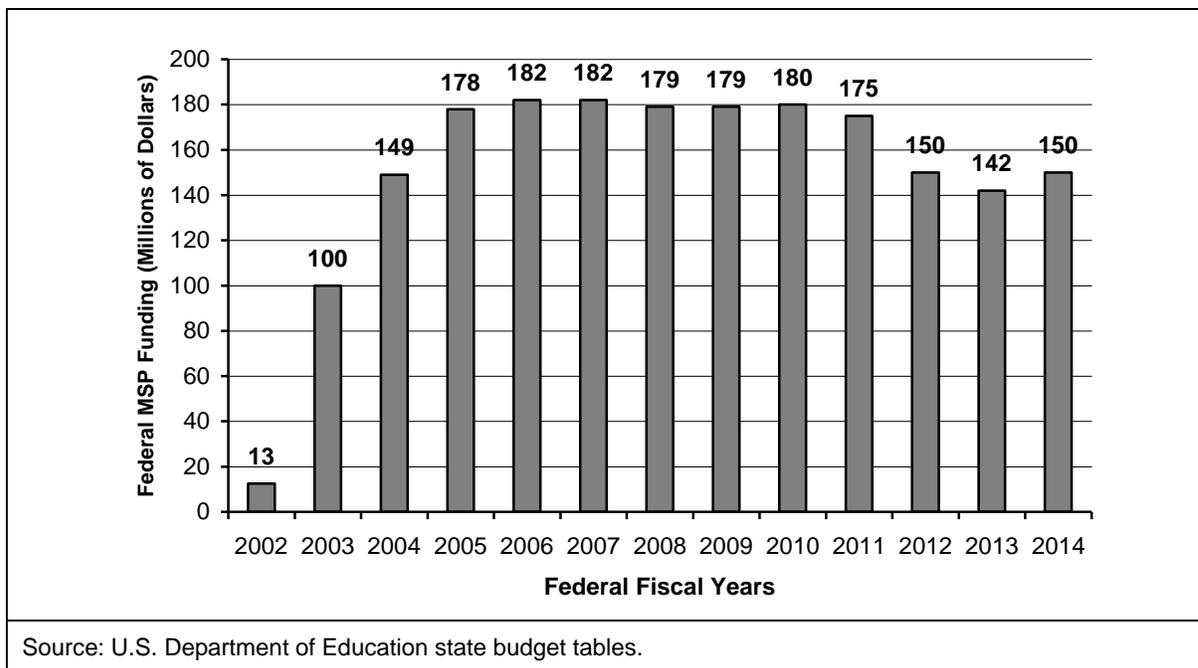
Implemented under the *No Child Left Behind Act of 2001, Title II, Part B*, the MSP program is strategically designed to improve the content and pedagogical knowledge of teachers and the academic performance of students in mathematics and science. The MSP program is a formula grant program to the states, with the size of individual state awards based on student population and poverty rates. The states then award grants on a competitive basis to local partnerships that are made up of, at a minimum, high-need schools or school districts⁶ and science, technology, engineering, and mathematics departments in IHEs. Other partners may include additional local education agencies, public or private schools, and businesses and non-profit or for-profit organizations.

⁶ The term “high-need” is not explicitly defined in the statute for the Mathematics and Science Partnership Program. Each state educational agency is responsible for conducting a needs assessment to determine the highest priority for these professional development funds and for defining high-need for its grant competition.

Exhibit 2 shows how federal support for the MSP program increased substantially from the program's inception in FY 2002 (\$12.5 million) to FY 2003 (\$100 million), when MSP became a state-administered formula grant program. Funding increased further between 2005 and 2011, during which time total funding for the program hovered around \$180 million annually. In the past three years, funding has decreased slightly.

In FY 2012, the period described in this report, states awarded \$150 million in funds to 488 local partnerships (sub-grants). State grants for FY 2012 ranged from approximately \$744,000 up to \$18 million with an average of \$2.7 million and a median of \$1.6 million. These grants provided professional development services to an estimated total of over 40,000 educators. Moreover, many projects trained teacher leaders, who then provided additional training to other teachers in their schools and districts.⁷

Exhibit 2: MSP Program Funding, Fiscal Years 2002–2014



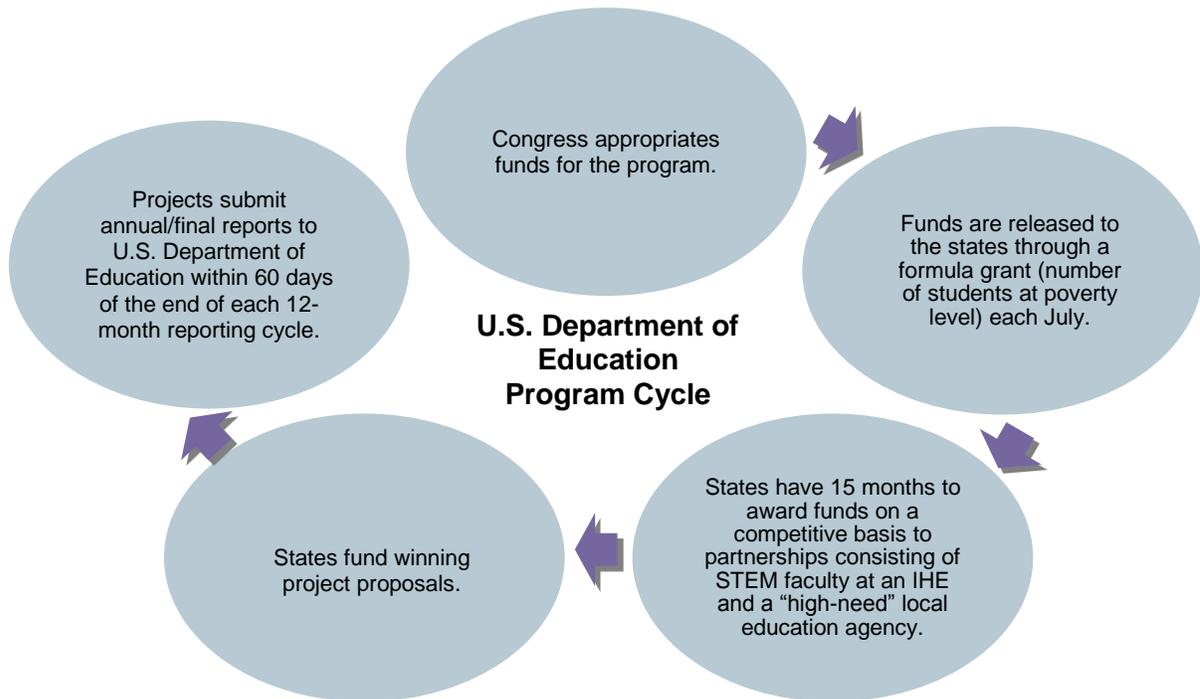
The administration of the MSP program involves an annual cycle of activities conducted at the federal, state, and local agency levels (see Exhibit 3). Each July, the Department of Education is charged with distributing MSP program funds to state education agencies for the upcoming fiscal year, based upon the number of children aged 5 through 17 years old in the state who live in families with incomes below the poverty line. In turn, states are required to run a competitive grant process to identify MSP projects and provide technical assistance to funded projects. Since FY 2003, all 50 states, the District of Columbia, and Puerto Rico have received MSP formula grants.⁸

⁷ Only teachers who received direct professional development through the MSP program are included in these numbers. Teachers who received training from teacher leaders trained through the MSP program are not included.

⁸ The American Virgin Islands, Guam, Northern Mariana Islands, and Samoa pool their MSP funds as part of their consolidated budget.

States have 15 months (through September 30 of the following year) to manage competitions and award their funds to projects (Exhibit 3). MSP sub-grants may be funded for up to three years. The law also requires all MSP projects to report annually to the U.S. Department of Education. Projects provide descriptive information and report progress toward meeting their goals in an on-line reporting instrument.

Exhibit 3: MSP Grant and Funding Cycle



Projects respond to both open-ended and closed-ended questions, and are required to report the following types of information in their APRs:

- Roles and responsibilities of MSP partners,
- Characteristics of MSP participants,
- Professional development models and content,
- Program evaluation design, and
- Evaluation findings and evidence of outcomes.

Report Overview and Analytic Approach

This report presents a summary of the data for projects funded in Performance Period 2012 (PP12).⁹ The findings presented in this report are primarily based on annual performance report (APR) data

⁹ Performance Period 2012 (PP12) refers to the period between October 1, 2012 and September 30, 2013. PP12 projects are those for which the majority of months of activities described in the Annual Performance Report take place in the 2012 fiscal year, between October 1, 2012 and September 30, 2013.

submitted by all MSP projects by February 28, 2014.¹⁰ Additionally, to examine trends in the MSP program over time, data from previous years are also included for some APR items. The report includes findings on a few selected APR items from previous periods beginning in PP06. However, for most items, trends are only examined over the past three years. Since there is substantial turnover in the set of projects included in the analyses for each year, the findings should not be thought of as longitudinal. Thus, we would not necessarily expect to see growth over time, as new projects are continually added to the program and other projects are ending.

The analyses were guided by five research questions (Exhibit 4). The first four research questions are addressed through the use of simple descriptive statistics, such as means and percentages from closed-ended questions in the APR. Additionally, to help illustrate the types of professional development activities offered, and the impact of the projects on teachers, students, and faculty, the open-ended item responses were examined, and examples are provided throughout the report as well as in a chapter on special topics relevant to MSPs. The fifth research question is addressed through the review of final-year MSP projects that reported using an experimental or quasi-experimental comparison-group design to assess their MSP programs.

Exhibit 4: Research Questions that Guide Analyses

RQ1	How are MSP projects implemented?
RQ2	Do MSP projects report using rigorous designs, such as experimental or quasi-experimental designs, for their evaluations?
RQ3	Do teachers that participate in the MSP program increase their scores on assessments of content knowledge?
RQ4	Do students in classrooms of teachers that participate in the MSP program score at the proficient level or above in state assessments of mathematics or science?
RQ5	Do MSP projects using an experimental or quasi-experimental design for their evaluations conduct their evaluations successfully and do they yield scientifically valid results?

¹⁰ These primarily included PP11 reports, but they also included some PP10 reports for which teacher and/or student data were not available in time to submit during the previous year.

Report Organization

The remainder of this report is organized into six chapters and three appendices, as follows:

Chapter 2: Characteristics of MSP Projects and Participants

Chapter 3: Professional Development Content and Activities

Chapter 4: MSP Evaluation Designs and Outcomes

Chapter 5: Highlights from PP11 MSP Projects with Rigorous Designs

Chapter 6: Summary and Conclusions

Appendix A: Review of Projects with Rigorous Designs

Appendix B: Criteria for Classifying Designs of MSP Evaluations

Appendix C: 2012 State MSP Appropriations

Chapters 2 and 3 describe how MSP projects were implemented. Chapter 4 describes the designs and outcomes projects reported. Chapter 5 presents highlights from PP12 MSP projects that implemented rigorous evaluations. Finally, Chapter 6 provides a summary of the findings and makes concluding comments.

Appendix A provides a review of the final evaluation designs of projects that reported using experimental or quasi-experimental designs; Appendix B contains the criteria used for classifying rigorous evaluation designs; and Appendix C presents a table with the 2012 MSP state appropriations.

Chapter 2: Characteristics of MSP Projects and Participants

This chapter describes the general characteristics of the MSP projects. It provides information on the sources and amounts of funding used by MSP projects, the types and number of partners involved in MSP projects, the number of teachers and students served by MSP projects, the characteristics of those teachers, and the methods of participant selection.

Sources and Amounts of Funding

The MSP program is a formula grant program to the states, with the size of individual state awards based on student population and poverty rates. In PP12, federal MSP resources totaling \$150 million were distributed through formula grants to all 50 states, the District of Columbia, Puerto Rico, and U.S. Island areas.¹¹ No state received less than one half of one percent of the total appropriation; MSP appropriations to individual states ranged from \$744,840 to \$17.9 million. See Appendix C for the specific MSP appropriation to each state.

With these funds, each state is responsible for administering a grant competition in which grants are made to partnerships to improve teacher knowledge in mathematics and science. Individual MSP awards ranged from \$20,000 to \$5.8 million with an average of \$344,643 and a median of \$213,456. As shown in Exhibit 5, over three-quarters of projects (78 to 86 percent) received an award of \$500,000 or less between PP06 and PP12. The size of awards in PP12 has continued the trend seen in recent years, with most projects receiving awards between \$100,000 and \$500,000, and fewer projects receiving either smaller or larger awards.

Exhibit 5: MSP Awards from State MSP Grants, Performance Periods 2006–2012

Project Budgets	Percent of Projects						
	PP06 (N=488)	PP07 (N=574)	PP08 (N=626)	PP09 (N=588)	PP10 (N= 566)	PP11 (N= 498)	PP12 (N= 487)
\$100,000 or less	17%	9%	13%	13%	12%	12%	16%
\$100,001 to \$200,000	37	43	38	37	36	34	32
\$200,001 to \$500,000	26	26	30	36	37	38	35
\$500,001 to \$1,000,000	15	18	17	13	12	14	13
1,000,001 or more	5	4	2	1	3	3	3

Source: Annual Performance Report item I.A.6

The non-response rate was 1 percent in PP06, <1 percent in PP07, 0 percent in PP08, <1 percent in PP09, <1 percent in PP10, 0 percent in PP11, and <1 percent in PP12

¹¹ The American Virgin Islands, Guam, the Northern Mariana Islands, and Samoa pool their MSP funds as part of their consolidated budget. They are not required to submit annual performance reports to the MSP Program, so their activities are not reflected in this report.

Some MSP projects supplemented their federal MSP funds with funds from other federal and non-federal sources. In PP12, 10 percent of projects reported receiving funds from other sources. These additional funds ranged from \$450 to \$445,000.

Organization and Partnerships

Each MSP grant has a lead organization that serves as the designated fiscal agent for the project. The lead organization is primarily responsible for distributing MSP funds, but often organizes and manages the project’s activities as well. The lead organization is typically either a local school district or an institution of higher education (IHE), as seen in Exhibit 6. In all but three years (PP09–PP11), over half of all projects (between 53 and 56 percent) had local school districts serve as fiscal agents, while approximately one-third of projects (31 to 37 percent) had IHEs fulfill this role. The remaining projects indicated that neither a local school district nor an IHE served as the lead organization. The other designated fiscal agents in PP12 for the projects primarily included regional organizations (9 percent) and non-profit organizations (3 percent).

Exhibit 6: Types of Lead Organizations, Performance Periods 2006–2012

Type of Lead Organization	Percent of Projects						
	PP06 (N=487)	PP07 (N=575)	PP08 (N=626)	PP09 (N=590)	PP10 (N=566)	PP11 (N=499)	PP12 (N=488)
Local school district	53%	56%	50%	47%	44%	45%	53%
Institution of higher education (IHE)	31	31	37	35	39	40	34
Non-profits, regional educational agencies, or other organizations	16	13	13	18	18	15	13

Source: Annual Performance Report item I.B.3
 The non-response rate was 1 percent in PP06, 0 percent in PP07, 0 percent in PP08, 0 percent in PP09, 0 percent in PP10, 0 percent in PP11, and 0 percent in PP12.

The MSP program establishes local partnerships that include: 1) a science,¹² technology, engineering and/or mathematics department of an IHE and 2) a high-need local education agency. In addition, MSP projects may incorporate other types of partners such as: education departments from IHEs; additional local education agencies including public charter schools, public or private elementary or secondary schools, and school consortia; and businesses and non-profit or for-profit organizations that have a proven capacity to effectively improve the knowledge of mathematics and science teachers. MSP projects reporting in PP12 had a median of 6 partner organizations, with the number of partners ranging from 1 to 690. Approximately half of all partner local education agencies (52 percent) were considered to be high-need.

¹² Computer science is included with science departments.

In PP12, 3,169 IHE faculty members, working in a variety of disciplines, were involved with MSP projects. As shown in Exhibit 7, half or more of all projects included faculty from science (57 percent) or mathematics (66 percent) departments; 16 percent of projects included faculty from engineering departments; and 14 percent of projects included faculty from technology departments. Additionally, nearly two-thirds of the projects (64 percent) reported working with faculty members from education departments, and 7 percent of projects included faculty from “other” departments, such as economics, psychology, and political science, as well as individuals associated with IHEs in a capacity other than teaching faculty, such as deans, administrators, district services, K–12 outreach staff, and consultants. On average, 7 IHE faculty members participated per project, from multiple disciplines.

Exhibit 7: Disciplinary Affiliation of IHE Faculty Participating in MSP, Performance Period 2012

Discipline	Percent of Projects (N=488)	Average Number per Project	Total Number Participating in MSP (Sum = 3169)
Science	57%	4	1,077
Mathematics	66	3	948
Engineering	16	2	180
Education	64	3	807
Technology	14	2	100
Other	7	2	57

Source: Annual Performance Report items IV.A.1– 5
 Percents total more than 100 percent because respondents could check more than one category.
 The non-response rate was 0 percent.
¹Computer science is included together with science.

MSP projects classified their stage of implementation into one of three stages: (1) *new*, defined as conducting start-up tasks such as planning activities, formalizing partnerships, and implementing the professional development model for the first time; (2) *developing*, defined as revising, enhancing, or continuing to develop their professional development model; and (3) *fully developed*, defined as having all components of the project’s planned model fully operational. Exhibit 8 shows that in PP12, more projects reported being fully developed or developing than new (61 percent, 24 percent, and 15 percent of projects, respectively). This trend is in keeping with a continuing increase in the proportion of projects that consider their implementation to be fully developed.

Exhibit 8: Projects’ Stage of Implementation, Performance Periods 2010–2012

Stage of Implementation	Percent of Projects		
	PP10 (N=566)	PP11 (N=497)	PP12 (N=488)
Stage 1: New	16%	16%	15%
Stage 2: Developing	34	28	24
Stage 3: Fully Developed	50	56	61

Source: Annual Performance Report item VII.C
 The non-response rate was 0 percent in PP10, <1 percent in PP11, and 0 percent in PP12.

Number of Participants Served by MSP

The central purpose of the MSP program is to provide professional development to teachers in order to increase their mathematics and/or science content knowledge and their pedagogical skills. The underlying logic is that with deeper knowledge of the subject matter and understanding of effective instructional strategies, teachers will be better able to impact their students' achievement in mathematics and science. To accomplish this goal, MSP projects work with a variety of teachers, across grades K through 12. Additionally, the program aims to increase the support structures in place for these teachers by training teacher leaders, coaches, and paraprofessionals, and by promoting the instructional leadership of administrators.

MSP projects reported serving over 40,000 participants in PP12, including elementary, middle, and high school teachers, coaches, paraprofessionals, and administrators (Exhibit 9). Sixty percent of educators reported working in high-needs schools. The total number of participants represents a slight decrease in the number of participants served from previous years. The median number of participants served per project has increased slightly over the past few years.¹³ The number of participants reported by individual projects varied widely, ranging from a low of 6 participants to a high of 1,382. Nearly all projects (91 percent) worked with 200 participants or fewer. Over half of the projects (53 percent) reported serving 50 or fewer participants in PP12; 27 percent reported serving between 51 and 100 participants; and the remaining projects (20 percent) reported serving more than 100 participants.

Exhibit 9: Distribution and Statistics Regarding Total Number of Participants Served by MSP Projects, Performance Periods 2010–2012

Number of Participants Served	PP10 (N=566)	PP11 (N=499)	PP12 (N=481)
Total number served by MSP projects	43,755	43,146	40,052
Median number served per project	41	45	47
Minimum number served per project	5	7	6
Maximum number served per project	1,200	1,781	1,382
25 or fewer	20%	20%	23%
26–50	39	38	31
51–100	24	25	27
101–200	10	10	11
201 or more	7	8	9
Source: Annual Performance Report items IV.C, IV.G.1			
The non-response rate was <1 percent in PP10, 0 percent in PP11, and 0 percent in PP12.			

¹³ A median of 47 means that half of reporting MSP projects served 47 or fewer participants, and half served more than 47 participants. The median is a more meaningful measure of the number of participants served by typical projects since the mean number of participants was heavily skewed by a few projects that reported serving more than 500 participants.

Methods of Selecting Participants

MSP projects are encouraged to identify and select schools and teachers for participation according to the level of need for professional development services in mathematics and science. Nearly three-quarters of projects (74 percent) indicated that the main goal of their MSP project was to improve individual teachers' content knowledge, while only 2 percent had the main goal of training teacher leaders who would in turn train other teachers (Exhibit 10). Twenty-two percent of projects reported that both goals were equally important, indicating that most projects that train teacher leaders also train individual teachers. An additional 3 percent of projects reported another type of main goal, such as sustaining and documenting their activities, embedding the use of technology in instruction, using formative assessments, or and training on reform-based pedagogy.

Exhibit 10: Main Goal of MSP Project, Performance Period 2012

Main Goal	Percent of Projects (N=488)
Improving teachers' content knowledge	74%
Training teacher leaders	2
Both improving teachers' content knowledge and training teacher leaders	22
Other	3
Source: Annual Performance Report item IV.B.1 The non-response rate was 0 percent.	

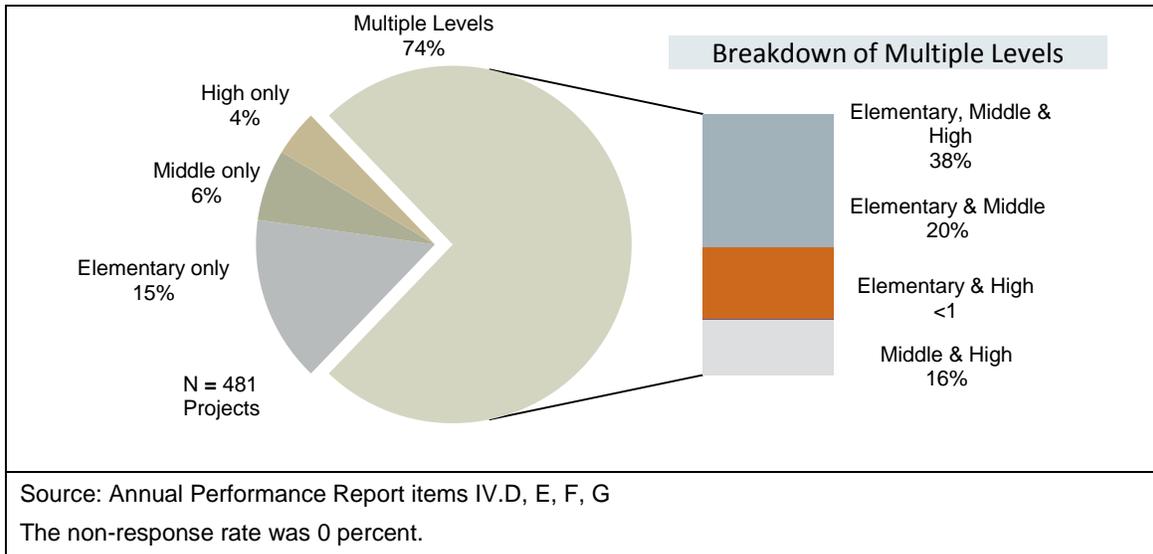
School Levels and Types of Participants Served

MSP projects are structured to address the professional development needs of educators at varying levels of the K–12 system. Projects may work with a group of participants drawn from a single school level (elementary, middle, or high school), participants from a combination of school levels, or participants from the entire K–12 spectrum. Overall, in PP12, nearly three-quarters of projects (74 percent) worked with participants from multiple school levels, while 26 percent of projects targeted a single school level.

As shown in Exhibit 11, 15 percent of all MSP projects in PP12 targeted the elementary school level only, 6 percent targeted the middle school level only, and 4 percent targeted the high school level only. Among projects that targeted multiple school levels, 38 percent of projects targeted participants at all school levels; 20 percent targeted elementary and middle school participants; 16 percent targeted middle and high school participants; and less than 1 percent targeted elementary and high school participants.

MSP participants were distributed across school levels in PP12 as follows: 47 percent at the elementary level, 28 percent at the middle school level, and 25 percent at the high school level. This distribution has remained fairly stable over recent years.

Exhibit 11: School Levels of Participants Served, Performance Period 2012



The MSP projects serve a variety of educators at all school levels, including classroom teachers, administrators, and other school staff. Exhibit 12 examines the different types of educators participating in MSP projects and shows the total proportion of each participant type served, by school level.

The most commonly reported MSP participants, across all school levels, are “regular core content” teachers, defined as elementary school teachers who have regular classroom assignments, and middle and high school teachers with mathematics, science, or technology assignments. At each school level, 86 percent of teachers were regular core content teachers. Other types of MSP participants include:

- *Special education teachers*—teachers who teach or support children with special learning needs;
- *School administrators*—both principals and assistant principals;
- *Mathematics and science coaches*—specialists who provide direct one-on-one coaching to students, and specialists who work with teachers to model instruction, conduct classroom observations, and provide personalized feedback and support;
- *Teachers of English language learners (ELL)*—teachers who offer support to students whose primary language is a language other than English;
- *Gifted and talented/Advanced Placement (AP)/International Baccalaureate (IB) teachers*—teachers who specialize in working with gifted students who need additional challenge; and
- *Paraprofessionals*—staff, often referred to as aides, who are not licensed to teach, but who perform many educational duties, both individually with students and organizationally in the classroom.

The next two largest groups of MSP participants across school levels were special education teachers (between 5 and 6 percent) and school administrators (between 3 and 4 percent).

Exhibit 12: Percent of Teachers and Other School Staff Among All MSP Participants Served, by School Level, Performance Period 2012

Participant Type	Percent of Teachers and Other School Staff Served		
	Elementary School (K–5) (N=18,635)	Middle School (6–8) (N=10,993)	High School (9–12) (N=9,825)
Regular core content	86%	86%	86%
Special education teachers	5	6	5
School administrators	3	3	4
Math coaches	2	1	1
Science coaches	<1	<1	<1
ELL	2	1	1
Gifted and talented / AP-IB	1	2	3
Paraprofessionals	<1	<1	<1

Source: Annual Performance Report items IV.D, E, F, G
 The non-response rate was <1 percent.
 Administrators who received professional development are not included in this exhibit.

In total, MSP projects reported reaching over 2.3 million students in PP12. Exhibit 13 shows the total number of students at each school level who were taught by MSP participants, as well as the median¹⁴ number of students reached by MSP participants.

Exhibit 13: Total Number of Students Taught by Participants in MSP Projects, Performance Period 2012

Number of Students Taught	Elementary School (N=351 Projects)	Middle School (N=384 Projects)	High School (N=275 Projects)
Total number taught by MSP participants	575,706	853,610	883,560
Median number taught per project	919	1,159	1,095
Minimum number taught per project	20	6	4
Maximum number taught per project	40,890	45,252	94,362

Source: Annual Performance Report items IV.H
 The non-response rate was 3 percent.
 Projects could serve one or multiple school levels.

¹⁴ These data, similar to the data on number of teachers, have been skewed by the presence of several unusually large projects. Therefore, the median is used to illustrate the number of students reached by a typical MSP project.

Chapter 3: Professional Development Content and Activities

This chapter describes the professional development activities offered in MSP projects. First, it describes the specific mathematics and science content of the MSP professional development. Then it describes the models of professional development offered (i.e., whether the professional development was primarily offered through summer institutes with follow-up or whether it focused on academic-year activities) as well as the specific learning activities within those professional models.

Professional Development Content of MSP Projects

In their annual reports, projects indicated whether they provided mathematics and/or science content in their MSP professional development. They also identified the major topics within each discipline and the grade level of the teachers to whom each topic was taught. As shown in Exhibit 14, in PP12, 40 percent of projects focused on mathematics only, 24 percent focused on science only, and 36 percent focused on both mathematics and science. Although the trend of more projects focusing on mathematics than on science has remained fairly stable over time, in PP11 and PP12 the gap between the proportions of mathematics and science projects widened.

Exhibit 14: Content Focus of Professional Development, Performance Periods 2010–2012

Content Focus	PP10 Percent of Projects (N=565)	PP11 Percent of Projects (N=496)	PP12 Percent of Projects (N=480)
Mathematics only	38%	41%	40%
Science only	31	26	24
Mathematics and science	31	33	36

Source: Annual Performance Report items VI.A.1, VI.B.1
The non-response rate was 0 percent in PP10, <1 percent in PP11, and <1 percent in PP12.

MSP projects that provided professional development in both mathematics and science determined whether to integrate content delivery across the two subjects. Projects that used an integrated approach offered joint professional development opportunities on mathematics and science topics, while projects that did not integrate them taught mathematics and science courses separately, either contemporaneously or consecutively.

Mathematics Content and Processes

Almost every MSP project provided professional development in multiple content areas and processes, often focusing on topics relevant to the grade level of the participating teachers. Across MSP projects, these areas included: number and operations, algebra, geometry, measurement, probability and statistics, problem solving, reasoning and proof, and calculus. Exhibit 15 disaggregates these areas to show how often each topic was addressed across all projects; however, most projects covered more than one topic. In mathematics, problem solving was the most frequently addressed *topic* across all school levels (83 to 88 percent of projects), while number and operations was the most commonly addressed *content area* in elementary school (80 percent), and algebra was the most frequently addressed *content area* in middle and high school (74 percent and 73 percent, respectively). Calculus was the least frequently addressed topic across all school levels (4 to 22 percent).

Exhibit 15: Content Areas and Processes of Mathematics Professional Development Provided to Teachers, by School Level, Performance Period 2012

Mathematics Content and Processes	Elementary School Teachers Percent of Projects (N=250)	Middle School Teachers Percent of Projects (N=272)	High School Teachers Percent of Projects (N=203)
Problem solving	87%	88%	83%
Number and operations	80	70	56
Measurement	65	66	64
Algebra	58	74	73
Reasoning and proof	56	60	57
Geometry	49	59	61
Probability and statistics	44	54	59
Calculus	4	10	22
Other	18	17	19

Source: Annual Performance Report item VI.A.2

The total number of projects that provided professional development in mathematics content areas or processes in PP12 was 364. The non-response rate was 1 percent.

Percents total more than 100 percent because respondents could check more than one category. Projects could serve one or multiple school levels.

Science Content and Processes

As in mathematics, professional development in science was provided in topic areas relevant to the grade level of the participating teachers. Projects also focused on multiple content areas and processes in and across disciplines. Across MSP projects, these areas included: scientific inquiry, physical science/physics, chemistry, life science/biology, earth science, and technology. As shown in Exhibit 16, scientific inquiry was a commonly addressed *topic* among projects across all school levels that addressed science (92 to 98 percent of projects), since it can be used across content areas, and technology was addressed by approximately two-thirds of projects (64 to 71 percent). Physical science and/or physics were the most commonly addressed *content areas* (69 to 73 percent) across all school levels, and chemistry was the least frequently addressed topic (41 to 46 percent). Additionally, many projects (64 to 71 percent) across all school levels provided professional development in technology.

Exhibit 16: Content Areas and Processes of Science Professional Development Provided to Teachers, by School Level, Performance Period 2012

Science Content Areas and Processes	Elementary School Teachers Percent of Projects (N=213)	Middle School Teachers Percent of Projects (N=219)	High School Teachers Percent of Projects (N=156)
Scientific inquiry	98%	96%	92%
Technology	64	65	71
Physical science/Physics	71	73	69
Life science/Biology	61	63	60
Earth science	57	58	47
Chemistry	41	45	46
Other	15	18	20

Source: Annual Performance Report item VI.B.2

The total number of projects that provided professional development in science content areas or processes in PP12 was 288. The non-response rate was 1 percent.

Percents total more than 100 percent because respondents could check more than one category. Projects could serve one or multiple school levels.

Professional Development Models

MSP partnerships often focus their professional development activities around a “summer institute,” which is defined in MSP’s governing legislation as a model of professional development that provides intensive learning experiences over a minimum of a two-week period. Although improving teacher content knowledge directly through a summer institute with in-school follow-up is a common model of MSP professional development, most of the remaining projects also include shorter summer components in addition to academic-year activities. In the following sections, we describe a variety of professional development activities.

Projects with Summer Institutes

In PP12, just over half of MSP projects (53 percent) conducted a summer institute. These learning experiences include deep exploration of mathematics and science content. Projects that offer summer institutes typically conduct follow-up activities during the academic year, with the aim of enhancing or extending the knowledge gained by participants over the summer. Nearly all of the projects that offered summer institutes also conducted follow-up activities. As shown in Exhibit 17, in PP12, 50 percent of projects conducted summer institutes with academic-year follow-up activities, while only 3 percent conducted summer institutes without any academic-year follow-up activities.

Exhibit 17: Types of Professional Development Models, Performance Period 2012

Professional Development Model	Percent of Projects (N=481)
Summer institute only	3%
Summer institute with follow-up activities	50
Focus on academic-year activities	47
Source: Annual Performance Report item V.B The non-response rate was 0 percent.	

Projects Offering Other Models of Professional Development

The 47 percent of MSP projects that did not conduct a two-week summer institute in PP12 provided other types of professional development activities that primarily took place during the academic year, many of which also included a summer component. While some professional development may have taken place over the summer, these activities did not fit into the definition of “summer institute,” which requires a minimum of two weeks of professional development. Instead, they were likely to include shorter professional development sessions or workshops interspersed throughout the summer months as well as during the academic year. Among the projects that did not offer a two-week summer institute, nearly half (46 percent) reported offering between one and two weeks of professional development in the summer, and 30 percent offered less than one week of professional development in the summer. The remaining 24 percent of projects held all of their professional development activities during the academic year.

Hours of Professional Development Provided

Exhibit 18 shows the median number of hours of professional development¹⁵ provided by model type. Among projects that conducted summer institutes only, a median of 80 hours of professional development were provided, and projects that focused on academic-year activities provided a median of 68 hours. Projects that conducted summer institutes with follow-up activities provided a median of 97 hours. When the time spent during the summer was analyzed separately from academic-year activities, projects spent a median of 66 hours during the summer institute, and a median of 30 hours on follow-up activities.

¹⁵ Projects that provided a very high or very low level of professional development skewed the average (mean), so we present the median.

Exhibit 18: Median Hours of Professional Development, By Model Type, Performance Period 2012

Professional Development Model	Median Number of Hours
Summer institute only	80
Summer institute with follow-up activities:	97
<i>Summer institute portion</i>	66
<i>Follow-up activities portion</i>	30
Focus on academic-year activities	68
Source: Annual Performance Report item V.A.1, V.B(i).1, V.B(ii).1	
The non-response rate for each model was as follows: summer institutes only: 0 percent; summer institutes with follow-up: 8 percent; and focus on academic-year activities: 0 percent.	
Medians are calculated separately within each category. The medians for each type of follow-up do not sum to the median of the whole.	

Primary Academic Year Professional Development Activities

In addition to providing intensive professional development during the summer, MSP projects offered a wide range of other professional development activities during the academic year to participating teachers in PP12. Such activities were offered as follow-up to summer institutes, to supplement material and concepts learned in those institutes, or in lieu of summer institutes. In this section, we first present the prevalence of these additional activities; then we describe each type of professional development activity and provide examples cited by projects of how these primary activities were implemented. The examples help to provide a sense of the broad variety of activities in which projects are engaged.

Exhibit 19 summarizes the primary academic year professional development activities that projects reported. Overall, the most common form of academic-year professional development reported by MSP projects in PP12 was on-site professional development, which often takes place at or near the teachers' schools. This category includes activities such as recurring workshops, coaching, and mentoring, and was reported by 69 percent of projects that offered academic-year activities. The next most common form of academic year professional development reported was study groups, which was reported by 14 percent of projects. Other reported activities include coursework at universities (6 percent) and on-line course work/distance learning networks (3 percent), while the remaining 8 percent of projects reported that they offered professional development activities that did not fall into one of the previously mentioned categories, such as field experiences, mentoring, conference attendance, or workshops.

Exhibit 19: Primary Form of Academic Year Professional Development Activities, Performance Period 2012

Primary Focus of Professional Development Activities	Percent of Projects (N=462)
On-site activities during academic year	69%
Study groups	14
University courses	6
On-line coursework / distance learning networks	3
Other activities (including field experiences, mentoring)	8
Source: Annual Performance Report items V.B(ii), V.B(iii) The non-response rate was 1 percent.	

The following sections describe each of the professional development activities in more detail and provide specific examples of how individual projects reported implementing these activities.

On-Site Activities during Academic Year

As noted above, 69 percent of all MSP projects reported that they engaged in on-site professional development activities as their primary form of professional development during the academic year. Examples of these on-site activities include mathematics and science content instruction, exploration of math and science education content standards, curriculum mapping, lesson and curriculum development, classroom modeling and demonstration, and classroom observation with feedback.

Depending on the project and the activity, these sessions were conducted either with groups of teachers within or across grade levels, or one-on-one between individual teachers and mentors or coaches. Examples of the types of mentors or coaches reported by various projects include fellow teachers, district staff members, institution of higher education (IHE) faculty, graduate students, and professional development providers. Mentors and coaches can provide direct one-on-one coaching or work with teachers to model instruction, plan lessons, conduct classroom observations, and provide personalized feedback and support.

Study Groups

Fourteen percent of the projects reported that their primary form of professional development during the academic year was study groups. Teacher study groups, which are sometimes structured as professional learning communities (PLCs), provide opportunities for ongoing collaboration with colleagues. Some projects reported that teachers in these groups shared lesson plans and reflected on both their content knowledge and classroom practice. Teachers might work with same-grade peers to better understand math and science education content standards, or participate in vertical teaming where they work with colleagues at consecutive grade levels to better understand the learning progression embodied in the standards and/or the curriculum. Other teacher groups engaged in lesson study, a process in which teachers jointly plan, observe, analyze, and refine actual classroom lessons.

Content Course Work at a College or University

With the goal of enhancing teachers’ content knowledge, 6 percent of projects reported courses provided by a local college or university as their primary form of professional development. The courses were often intensive and were held in the evenings or on weekends during the academic year. Some courses were also condensed into a period of two to three full-time weeks in the summer. In

some cases, teachers earned undergraduate or graduate credit, and completing the courses helped teachers meet requirements for certification or highly qualified status.¹⁶

On-Line Coursework/Distance Learning Networks

In order to provide teachers with convenient access to content materials, some MSP projects offered on-line courses or course modules that teachers could access on demand during the summer or academic year, and distance learning networks that help projects reach out to geographically isolated teachers. Three percent of projects reported this as their primary form of academic year professional development.

An advantage of on-line programs is that they allow expanded access to professional development for teachers in rural areas and those who need the scheduling flexibility. Like other content activities offered by MSP projects, on-line courses usually focus on mathematics or science content but might also address issues related to teaching and learning, curriculum development, assessment, or other topics. A project's on-line course might also utilize software applications that support on-line communities such as Blackboard or WebCT, to encourage collaboration and communication among participants and facilitators.

Whereas the main function of on-line coursework activities is content delivery, distance learning networks focus on increasing collaboration and support among participants and MSP facilitators. Teachers who would otherwise have had to travel long distances to meet with their counterparts or with university faculty are able to form communities and/or mentoring relationships through the use of email, message boards, phone contact, videoconferencing, and other communication technologies. Examples of professional development offered by distance learning networks include mentoring and coaching, lesson plan exchanges, on-line study group discussions, and blogging.

Other Activities

Eight percent of MSP projects reported other activities as their primary form of academic-year professional development. The variation among these other activities demonstrates how projects accommodated the varied needs and circumstances of participating schools and teachers. Some examples of "other activities" included weekend workshops, field training, conferences, and developing and posting videos.

¹⁶ A "highly qualified" teacher must 1) hold a bachelor's degree; 2) have a full state certification or license; and 3) have demonstrated subject matter competence in each of the subject area(s) taught.

Chapter 4: MSP Evaluation Designs and Outcomes

This chapter describes the types of evaluators and evaluation designs used by MSP projects, the measures used in the evaluations, and teacher and student outcomes, which are used to assess the effectiveness of the MSP interventions.

Evaluation Designs

Every MSP project is required to design and implement an evaluation and accountability plan that allows for a rigorous assessment of its effectiveness. Projects are required to report on two aspects of their evaluation findings: 1) gains in teacher content knowledge based on pre- and post-testing; and 2) proficiency levels on state-level assessments of students of teachers who received professional development.¹⁷

As seen in Exhibit 20, approximately two-thirds of projects (68 percent) reported using an external evaluator in PP12. Using external evaluators—specialized staff from outside the partnership trained to conduct evaluations—allows projects to independently evaluate their work, and to receive help from these specialists in implementing the most rigorous designs feasible. Nearly one-third of projects (30 percent) also reported involving their own partnership staff in their evaluations. This might have included their school system’s research office or a university research department. In addition, 14 percent of projects reported that they received support from their state to participate in a statewide evaluation, placing their project in context with the rest of the MSP work being done in their state.

Exhibit 20: Types of Project Evaluators, Performance Period 2012

Type of Evaluator	Percent of Projects (N= 488)
External evaluator	68%
MSP partnership organization staff	30
Statewide evaluation	14
Other	7

Source: Annual Performance Report item VII.A
 The non-response rate was 0 percent.
 Percents total more than 100 percent because respondents could check more than one category.

Exhibit 21 presents the types of evaluation designs that projects reported using in PP12. Projects that used a combination of designs were instructed to report on the most rigorous design used in the project. Nearly half of projects (49 percent) reported using an experimental or quasi-experimental design. Less than 1 percent of projects reported that they implemented an experimental design, which is the most rigorous research design for testing the impact of an intervention, wherein schools, teachers, or students are randomly assigned to treatment or control groups. Nearly half of the projects (48 percent) reported using a quasi-experimental, or comparison group design to compare the effects of the MSP program on participating teachers and/or their students to comparison, non-participating teachers and/or students. Nearly one-third of projects (30 percent) used a matched comparison group

¹⁷ Since not all teachers receive professional development each year, teachers are only required to be tested at least once during the life of the project. Additionally, student proficiency is only required to be reported in the periods following intensive professional development of teachers.

design, which attempts to show causality by demonstrating equivalence between groups at baseline or adjusting for any initial differences between groups, and 19 percent of projects reported using a non-matched comparison group.

The remaining 51 percent of projects reported using a less rigorous design type. Thirty-six percent of projects reported using pre- and post-tests to assess the gains of the teachers served by MSP as their most rigorous aspect of their evaluation. Twelve percent of projects reported using qualitative methods only, and 3 percent of projects reported using “other” methods, which included a mix of quantitative and qualitative methods.

Exhibit 21: Types of Evaluation Designs Used by Projects, Performance Period 2012

Evaluation Design	Percent of Projects (N=484)
Random assignment design (experimental)	<1%
Quasi-experimental design	48
<i>Matched comparison groups</i>	30
<i>Non-matched comparison groups</i>	19
One-group design	36
Qualitative / descriptive design	12
Other	3
Source: Annual Performance Report item VII.B The non-response rate was 1 percent.	

Measures Used in Evaluations

MSP projects use a variety of instruments to assess teacher knowledge, student achievement, and/or the extent to which teachers apply the lessons from the MSP professional development to their classroom instruction. Below, we discuss the measures that projects used to assess these outcomes.

Measures of Teacher Knowledge

All projects are required to administer pre- and post-tests during the year(s) in which their teachers receive intensive professional development. Projects used the MSP Program’s “Teacher Content Knowledge” (MSP TCK) macro to determine the number of teachers with significant gains in teacher content knowledge.¹⁸ Exhibit 22 presents the types of assessments used to measure teachers’ content knowledge in mathematics and in science and the types of assessments used to assess teachers’ classroom practices.

Standardized tests were the most frequently reported type of assessment utilized to assess teachers’ content knowledge both in mathematics (67 percent) and in science (48 percent). The next most frequently reported type of assessment for both mathematics and science was locally developed assessments that were not tested for validity (24 percent and 32 percent, respectively), followed by locally developed assessments with evidence of validity and reliability (10 percent and 21 percent,

¹⁸ The macro uses a statistical test called a dependent t-test (for 30 or more respondents) or the Wilcoxon signed ranks test (for less than 30 *but at least 6* respondents) to calculate, with 85 percent certainty, the number of teachers who showed substantive gains on content knowledge tests.

respectively. The remaining projects used self-report by teachers to assess their content knowledge, or other types of tests.

Exhibit 22: Types of Assessments Utilized to Assess Teacher Outcomes, Performance Period 2012

Assessment Type	Percent of Projects		
	Mathematics Content Knowledge (N=290)	Science Content Knowledge (N=221)	Classroom Practices and Beliefs (N=223)
Standardized test	67%	48%	43%
Local test, valid & reliable	10	21	12
Local test, not valid & reliable	24	32	15
Surveys or ratings	3	6	56
Other type of test	12	5	23

Source: Annual Performance Report item VII.D.1
 Percents total more than 100 percent because respondents could check more than one category.
 Only projects that provided professional development in each area and subsequently assessed those teachers responded to this question.

Among projects that measured classroom practices and beliefs, over half of projects (56 percent) reported using surveys or ratings by teachers, students, or other MSP participants. As seen in Exhibit 23, the most commonly reported existing assessments used to measure classroom practices and beliefs were the *Survey of Teacher Attitudes and Beliefs* (56 percent of projects), the *Teacher Efficacy Belief Instrument* (16 percent), the *Reformed Teaching Observation Protocol (RTOP)* (13 percent), and the *Surveys of Enacted Curriculum* (11 percent).

Exhibit 23: Assessments Utilized to Assess Teachers in Classroom Practices and Beliefs, Performance Period 2012

Classroom Practices and Beliefs Assessment Measure	Percent of Projects Utilizing this Assessment (N=225)
Survey of Teacher Attitudes and Beliefs	56%
Teacher Efficacy Belief Instrument	16
Reformed Teaching Observation Protocol (RTOP)	13
Surveys of Enacted Curriculum (SEC)	11
Inside the Classroom Observation Protocol	5
Oregon Mathematics Leadership Institute (OMLI)	0
Praxis 3	0
Other Assessment	51

Source: Annual Performance Report item VII.D
 Percents total more than 100 percent because respondents could select more than one measure.
 Only projects that provided professional development in this area and subsequently assessed those teachers responded to this question.

Exhibits 24 and 25 present the assessments projects used to measure teacher content knowledge in mathematics and science, respectively. Note that projects could have reported using more than one assessment instrument. The two most commonly reported assessments used for assessing mathematical content knowledge were *Learning Mathematics for Teaching (LMT)* (37 percent of projects) and *Diagnostic Mathematics Assessments for Middle School Teachers* (19 percent). For measuring content knowledge in science, the two most commonly reported assessments were *MOSART: Misconception Oriented Standards-Based Assessment* (21 percent) and *Diagnostic Teacher Assessments in Mathematics and Science (DTAMS)* (17 percent).

Exhibit 24: Assessments Utilized to Assess Teachers in Mathematics, Performance Period 2012

Mathematics Assessment Instrument	Percent of Projects Utilizing this Assessment (N=290)
Learning Mathematics for Teaching (LMT)	37%
Diagnostic Mathematics Assessments for Middle School Teachers	19
State Teacher Assessment	6
PRAXIS II	2
Knowledge of Algebra for Teaching	1
Precalculus Concept Assessment (PCA)	1
Praxis1	0
Other Assessment	51

Source: Annual Performance Report item VII.D
 Percents total more than 100 percent because respondents could select more than one assessment.
 Only projects that provided professional development in this area and subsequently assessed those teachers responded to this question.

Exhibit 25: Assessments Utilized to Assess Teachers in Science, Performance Period 2012

Science Assessment Instrument	Percent of Projects Utilizing this Assessment (N=221)
MOSART: Misconception Oriented Standards-Based Assessment	21%
Diagnostic Teacher Assessments in Mathematics and Science (DTAMS)	17
State Teacher Assessment	3
Assessing Teacher Learning about Science Teaching (ATLAST):	2
PRAXIS II	2
Force Concept Inventory	1
Other Assessment	64

Source: Annual Performance Report item VII.D
 Percents total more than 100 percent because respondents could select more than one assessment.
 Only projects that provided professional development in this area and subsequently assessed those teachers responded to this question.

Measures of Classroom Instruction

MSP projects also measured the extent to which teachers applied lessons from their MSP professional development to their classroom instruction. As shown in Exhibit 26, in PP12 84 percent of projects used questionnaires or other forms of self-reporting by teachers, and 69 percent of projects engaged in direct classroom observation to assess participants’ understanding and use of the content and strategies learned during MSP activities. The classroom observations can provide more objective, performance-based assessments of teacher classroom practices, while the questionnaires and other forms of self-reporting can provide valuable insights into teachers’ opinions about how their MSP experience improved their teaching methods.

Projects reported other approaches to measuring classroom instruction as well, some of which were used in conjunction with classroom observation or questionnaires. Forty percent reported conducting interviews or focus groups. Over a quarter (27 percent) of projects reported using lesson plan analysis. Fourteen percent of projects reported reviewing journals in which participants tracked lesson plans and reflected on classroom practice, and 12 percent reported videotaping lessons. Nine percent reported using “other” assessment methods, which included examining student assessment data as well as various other types of teacher self-reporting.

Exhibit 26: Methods of Evaluating the Application of MSP Professional Development to Classroom Instruction, Performance Period 2012

Measures	Percent of Projects (N=481)
Questionnaire/Self-report	84%
Classroom observation	69
Interviews/Focus groups	40
Lesson plan analysis	27
Journals	14
Videotaping	12
Blogs	6
Other	9

Source: Annual Performance Report item VII.E
 Percents total more than 100 percent because respondents could check more than one category.
 The non-response rate was <1 percent.

Evaluation Findings

As part of their evaluations, MSP projects are required to assess changes in teachers’ content knowledge in mathematics and/or science during the years in which they receive intensive professional development. Projects reported the number of MSP teachers who significantly increased their content knowledge in mathematics and/or science topics on project pre- and post-assessments.

Teacher Outcomes

Exhibit 27 presents data on the number of teachers participating in professional development courses in mathematics and science and the proportion who had pre- and post-assessment data available in

each of the past three performance periods.¹⁹ In mathematics, 23,546 teachers reported receiving professional development courses in PP12, and 56 percent of these teachers had assessment data available for the period. In science, 16,167 teachers reported receiving professional development courses in PP12, and 62 percent of these had assessment data available for the period.

Exhibit 27: Number of Teachers Served and Percent of Teachers Assessed, Performance Periods 2010–2012

Content Area	Total Number of Teachers Served			Percent of Teachers with Content Assessments (Pre-Post)		
	PP10	PP11	PP12	PP10	PP11	PP12
Mathematics	25,344	23,807	23,546	53%	45%	56%
Science	19,562	16,042	16,167	60	58	62

Source: Annual Performance Report items VIII.A.1, 2, 4, 5
Individual teachers who received professional development in both mathematics and science may be included in the number of both science and math teachers.

Exhibit 28 presents data for those teachers who were assessed for gains in content knowledge. Among the teachers assessed in PP12, 63 percent showed significant gains in mathematics content knowledge and 67 percent showed significant gains in science content knowledge. Teacher gains in math have remained between 61 to 65 percent for the past three years, while teacher gains in science have been decreasing over the same period.²⁰ As discussed above, 67 percent of projects used standardized tests to assess teacher content knowledge in math, and 48 percent used standardized tests to assess content knowledge in science. These tests may or may not be well aligned to the content teachers are being taught.

Exhibit 28: Percent of Teachers with Significant Gains in Content Knowledge, Among Teachers with Pre-Post Content Assessments, Performance Periods 2010–2012

Content Area	PP10	PP11	PP12
Mathematics	65%	61%	63%
Science	74	69	67

Source: Annual Performance Report items VIII.A.2, 3, 5, 6

Student Outcomes

Projects also reported the number of students served, assessed, and scoring at the proficient level or above in state assessments of both mathematics and science. As shown in Exhibit 29, over 1 million students in PP12 were taught by teachers who received professional development in mathematics, and nearly 800,000 students were taught by teachers who received professional development in science.

¹⁹ Projects are required to administer pre- and post-tests to each teacher who received professional development at least once during the course of the grant. MSP grants are up to three years long.

²⁰ Given that different teachers may be tested each year across varying tests, we would not necessarily expect an increasing trend in teacher gains.

Exhibit 29: Number of Students Served and Percent of Students Assessed, Performance Periods 2010–2012

Content Area	Total number of students taught by MSP teachers			Percent of students with content assessments		
	PP10	PP11	PP12	PP10	PP11	PP12
Mathematics	1,280,438	1,407,724	1,038,381	64%	50%	50%
Science	903,788	814,751	787,007	39	29	29

Source: Annual Performance Report items VIII.B. 1, 2, 5, 6
 Students who are taught by teachers receiving professional development in math and science may be double counted.

State assessment data were reported for 50 percent of students in mathematics and for 29 percent of students in science. The proportion of students being assessed at the proficient level or above in science remained relatively stable over the past three years, but decreased substantially in math in PP12 relative to previous years (see Exhibit 30). This may be due to the changes in statewide assessments being used in some states. In PP12, in mathematics, 55 percent of students scored at the proficient level or above, and in science, 69 percent of students scored at the proficient level or above.²¹

Exhibit 30: Percent of Students Taught by MSP Teachers Scoring at Proficient Level or Above, Performance Periods 2010–2012

Content Area	Proficient Level or Above		
	PP10	PP11	PP12
Mathematics	65%	64%	55%
Science	67	67	69

Source: Annual Performance Report items VIII.B. 2, 3, 4, 6, 7, 8
 In PP10 the non-response rates were 8 percent in mathematics and 7 percent in science; in PP11 the non-response rates were 10 percent in mathematics and science; and in PP12 the non-response rates were 10 percent in mathematics and science.

²¹ Numbers were aggregated across all grade levels and schools.

Chapter 5: Highlights from MSP Projects with Rigorous Designs

Among the 190 MSP projects that reported that PP12 was their final year, 21 projects met the MSP standards for rigorous evaluation designs in demonstrating the impact of their programs on teachers and students. In this chapter, we provide highlights from these 21 projects. Appendix A outlines the review process for selecting this set of projects, according to the criteria of rigorous evaluation design. Appendix B describes the criteria used to determine rigor of design. By reviewing the findings from these rigorous evaluations, we can gain insight into what aspects of professional development are associated with improvements in teacher content knowledge, student achievement, and/or teacher practices.

Although most of these projects that met the criteria evaluated multiple outcomes within their final year reports, only those aspects of their research that were conducted in a rigorous manner and pertain to the potential *impact* of MSP programs on teacher content knowledge, teacher practices, or student achievement are included in this chapter.

For each project with an evaluation that met the criteria for rigorous design, we provide information about its background, goals and professional development. The summaries of the projects' efforts and achievements that follow are based upon information included in the projects' evaluation reports and their PP12 APRs. Exhibit 31 provides information about each MSP project that met the criteria for evaluation rigor. Below we provide a brief overview of key findings.

Key Findings

Participants Trained

- The majority of projects (14) were designed for elementary and middle school teachers. Two projects were designed for both middle and high school teachers, and an additional three projects were designed for teachers across all school levels. Finally, two projects were designed for high school teachers exclusively.
- One project worked additionally with special education teachers and another included school administrators in the professional development workshops.

Professional Development Initiatives

- Ten projects provided professional development in math exclusively, while five focused on science exclusively. Six projects provided professional development in both math and science.
- The majority of projects (20) provided at least one summer component (a workshop or institute); 13 of these projects provided an intensive summer institute of at least 60 hours. One of these projects included the experience of working alongside a practicing scientist in his or her laboratory for a 10-week period over the summer.
- Seventeen projects provided follow-up activities in the form of weekend sessions, individualized and small-group coaching, on-line training and communities, workshops, or other training; two projects provided follow-up classroom observations and feedback from university faculty during the academic year.

- Five projects encouraged the development of leadership skills and “teacher leaders.”

Research Designs Used

- Twenty of the projects successfully employed quasi-experimental study designs (QEDs); one project used a randomized control trial (RCT).
- Among the 21 projects that successfully studied their program’s impact, 14 projects examined impacts on teacher content knowledge, 4 projects examined impacts on classroom practices, and 14 examined impacts on student achievement.
- Nine of the evaluations found positive impacts of the MSP on teacher content knowledge, two found positive impacts on teacher classroom practice, and seven found positive impacts on student achievement.

Assessments Used

- A variety of assessments were used to evaluate the 14 projects that successfully studied their programs’ impacts on *teacher content knowledge*. Five projects used the Diagnostic Teacher Assessments in Mathematics and Science (DTAMS) assessment to evaluate impact. The other nine projects that used the following assessments: the Misconceptions-Oriented Standards-Based Assessment Resources for Teachers (MOSART) test, the Study of Instructional Improvement (SII), the Intel Math Content test, and project-developed assessments.
- The four projects with positive findings in *classroom practices* used the Reformed Teaching Observation Protocol (RTOP), the Survey of Enacted Curriculum (SEC), and project-developed assessments to evaluate impact.
- Among the 14 projects that successfully studied their program’s impacts on *student achievement*, 10 used state or district standardized test questions to measure achievement. The other four projects used the ThinkLink testing system, the Measures of Academic Progress (MAP), and project-developed assessments to evaluate impact.

Exhibit 31: MSP Projects with Rigorous Evaluation Designs

MSP Project	State	Participants	Content Area	Professional Development	Design of Evaluation(s)	Evaluations with Positive Findings
<i>Project SUCCESS</i>	CA	63 3rd–8th grade math teachers	Math	33-hour summer workshop; 31-hour workshop during the academic year; professional learning days; 8 hours of coaching	QED	Student achievement
<i>Collaboration for Leadership and Improvement in Math Education (CLIME) Project</i>	CA	60 elementary school teachers	Math	60-hour summer institute; 24 hours of follow-up training (3 school release days)	QED	Teacher content knowledge, student achievement
<i>Stanislaus County Math Partnership</i>	CA	92 5th–6th grade math teachers	Math	40-hour summer workshop; 32 hours of lesson study; 12 hours of individual instructional coaching	QED	None
<i>Downey Opportunities in Mathematics Partnership (DO Math)</i>	CA	45 3rd grade through Algebra I teachers	Math	60-hour summer institute; 24 hours of follow-up activities during the academic year, including participation in professional learning communities	QED	None
<i>Addressing Learning Gaps in Education by Raising Algebraic Understanding and Increasing Content Knowledge (ALGEBRAIC) Project</i>	CA	80 3rd grade through Algebra I teachers	Math	Five 8-hour summer workshops; 20 hours of mini-institutes	QED	Student achievement
<i>Inventing, Designing, Engineering, Activities in Science (IDEAS) Project</i>	CA	70 3rd–8th grade science teachers	Science	44-hour summer workshop; 40 hours of teaching learning collaboratives and after-school sessions during the academic year	QED	Teacher content knowledge
<i>Inquiry Core for 3rd, 4th and 5th</i>	CA	61 3rd–5th grade science teachers	Science	Three grade-specific summer sessions totaling 60 hours; 24 hours of follow-up during the academic year	QED	Teacher content knowledge
<i>Project DELTA (Developing Educators Learning to Teach Algebraically)</i>	CA	73 3rd–8th grade math teachers	Math	64-hour summer institute over 8 days; 48 hours of follow-up during the academic year, including two Saturday sessions	QED	Student achievement
<i>Algebra I Model Curriculum Field Study</i>	CT	55 Algebra I teachers	Math	Two summer institutes consisting of 8 units; two Saturday morning sessions; optional virtual discussion group led by a teacher mentor	QED	Student achievement

Mathematics and Science Partnerships: Summary of Performance Period 2012 Annual Reports

MSP Project	State	Participants	Content Area	Professional Development	Design of Evaluation(s)	Evaluations with Positive Findings
<i>Math and Science Endorsement Cohorts</i>	GA	19 elementary science teachers, 11 elementary math teachers	Math and Science	Four 50-hour content-focused courses which included participation in on-site training, field work, on-line discussions, and classroom observations	QED	None
<i>Graduate / Bradley University's Environmental Science Education Partnership Program</i>	IL	15 middle and high school teachers	Science	33 hours of graduate courses; 10-week summer experience with working scientist	QED	Teacher content knowledge
<i>WIP 3 ROE 45 MRI Math Connections II</i>	IL	51 elementary, middle school, special education, and secondary teachers	Math	80 hours of sessions over 8 units; summer institute; 4 days of follow-up training; participation in district-based team meetings	QED	Teacher content knowledge
<i>Promoting Teacher Quality and Student Achievement in Science</i>	NC	127 elementary and middle school teachers	Science	2-week summer institute; four Saturday follow-up sessions during the academic year; science leadership team participation	QED	Teacher content knowledge
<i>Assessing Core Content and Ensuring Success in Science (ACCESS)</i>	NC	150 Kindergarten–5th grade teachers	Science	2-week summer institute; two follow-up days during the academic year; professional learning communities; workshops for school administrators	QED	Teacher classroom practice
<i>Curriculum Topic Study to Enhance Achievement in Mathematics and Science (C-TEAMS)</i>	NJ	30 math and 36 science teachers in elementary through high school	Math and Science	70-hour summer institute; three workshops during the academic year; professional learning community; teacher exchange program	QED	Student achievement
<i>Greater Allentown Math Science Partnership</i>	PA	90 1st–8th grade teachers	Math and Science	80-hour summer institute; 32 hours of professional learning community activities; 8 hours of on-line professional learning community work; optional individualized classroom and curriculum coaching	RCT	Teacher content knowledge
<i>School District of Philadelphia Math and Science Partnership</i>	PA	20 high school teachers	Math and Science	Two 40-hour summer graduate-level math courses; optional additional math courses; Saturday science workshops and fellowship meetings	QED	Teacher content knowledge

MSP Project	State	Participants	Content Area	Professional Development	Design of Evaluation(s)	Evaluations with Positive Findings
<i>Project ARRMS: Achieving Rigor and Relevance in Math and Science</i>	PA	53 secondary math and science teachers	Math and Science	80-hour summer institute over 2 weeks; 24 follow-up hours of individualized and small-group instructional coaching; virtual discussions in on-line community	QED	Teacher classroom practices
<i>Building the Foundation to Algebra in Grades 3–8</i>	PA	55 3rd–8th grade teachers	Math	80-hour summer institute over 2 weeks; three follow-up days during the academic year; classroom observations	QED	Teacher content knowledge
<i>Laurens County STEM Teacher Development Program</i>	SC	60 3rd–12th grade math and science teachers	Math and Science	100-hour summer institute; ongoing professional development during the academic year; optional graduate STEM courses	QED	Student achievement
<i>Raising Achievement in Mathematics through Fostering Algebraic Thinking (RAM t-FAT)</i>	WI	30 middle and high school math teachers	Math	80-hour summer institute; mentoring and peer coaching; on-line community	QED	None
Sources: Performance Period 2012 APRs and Evaluation Reports						

Project SUCCESS

State (APR ID): California (CA100179)

Partners: 11 school districts, and mathematics and teacher education departments in four institutions of higher education: California State University Sacramento, Sierra Community College, National University, and Brandman University

Project Director: Chrissy Poulsen

Number of Participants: 63 3rd–8th grade mathematics teachers

Background: The goal of Project SUCCESS was to empower 3rd grade pre-algebra students with the tools and understandings which are critical to the development of their algebraic thinking by providing teachers with professional learning experiences designed to improve their mathematical and pedagogical content knowledge to address the specific needs of all students. This was hoped to reverse the decline in student proficiency in algebra and ultimately prepare students to enter into higher levels of mathematics.

Description of Professional Development: The program included included 33 hours of professional development over the summer and 31 hours during the academic year covering fractions, integer arithmetic, word problems, and variables. Additionally, professional learning days were provided which were facilitated by IHE partners, math consultants, and coaches. The professional learning days emphasized pedagogy (including lesson planning, developing assessments, and analyzing student errors) and analysis of student and teaching artifacts. Finally, teachers co-taught lessons with project staff and received eight hours of coaching including video recording in the classroom and in a professional learning community.

Before SUCCESS, the only way I knew how to teach math was to use the textbooks, give the kids formulas, and algorithms, and hope for the best. I dreaded problem solving. Now, after SUCCESS, I don't need a textbook to teach a concept and rather than giving a list of steps, I use "word problems" to teach each concept. I know how successful my students can be and that they don't have to all "get it" on the first example. The learning happens in the struggle!

—SUCCESS teacher

Description of Evaluation with Rigorous Design: The evaluation of *student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Student Achievement

The evaluator assessed whether students who were taught by teachers participating in the SUCCESS project performed better on a math assessment than the comparison group students who were taught by teachers who did not participate in the project. Math knowledge was measured using an assessment developed for the project. One analysis was conducted for students in grades 3 and 4 and another for students in grades 5 and 6, with sample sizes ranging from 504 to 519 for the treatment group and 305 to 520 for the comparison group. The evaluator reported that for both analyses, after adjusting for baseline differences, students of the teachers in the SUCCESS group significantly outperformed the comparison group students on the post-test, although the effect size was too small to be educationally meaningful.

Collaboration for Leadership and Improvement in Math Education (CLIME) Project

State (APR ID): California (CA100181)

Partners: 3 School districts and 3 private schools in Northern Orange County, Chapman University, and California State University Fullerton

Project Director: Amy Edmundson

Number of Participants: 60 teachers from 26 elementary schools

Background: The goals of the CLIME project included increasing teachers' mathematical content knowledge, pedagogy skill, and problem-solving abilities and bolstering teacher's confidence, leadership aptitude, and data-driven decision making. CLIME also tried to improve IHE faculty's instructional abilities through increased awareness of and participation in teacher preparation courses.

Description of Professional Development: Teachers participated in a 60-hour Math Institute during the summer administered by university faculty and a professional development coordinator. The institute covered topics including base ten, fractions, algebra, and general mathematical reasoning. Teachers also received information about aligning their teaching with grade level standards, created and taught an elementary school math lesson, and used research to discuss the effect of these lessons. Teachers received three school release days (24 hours) for follow-up training about aligning to standards, grouping students, the pacing of lessons, and how to monitor students' learning progress. The professional development coordinator also observed the CLIME teachers while they taught.

At [one university], [they] are re-examining the tasks [they] engage pre-service teachers in so as to prepare them better for the things they will be asked to do when they become teachers. This is a direct result of the things [the faculty member has] observed with CLIME teachers.

—CLIME faculty member

Description of Evaluation with Rigorous Design: The evaluations of *teacher content knowledge and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluator assessed whether teachers in grades 3–6 participating in the CLIME project showed greater increase in math content knowledge than comparison group teachers who did not participate in the project. This was examined in each of the three years of the program and for multiple measures of teacher content knowledge. Teacher content knowledge was assessed using subscales from the Diagnostic Teacher Assessments in Mathematics and Science (DTAMS) assessment and items developed for the project. The sample sizes for these analyses ranged from 68 to 72 for the treatment group and 37 to 71 for the comparison group. The evaluator reported that in year 1, after adjusting for baseline differences, the two groups of teachers did not differ significantly on their gain scores for the DTAMS subscales. In year 2, teachers in the CLIME project made significantly greater gains on the pedagogical content subscale of the DTAMS assessment, but did not show any significant differences on any of the other subscales. However, in year 2 CLIME teachers did show significantly greater gains on the project-developed geometry assessment. In year 3, the CLIME group teachers made significantly greater increases than the comparison group teachers on several subscales of the DTAMS measurement.

Student Achievement

The evaluator assessed whether treatment group students who were taught by teachers participating in the CLIME project outperformed comparison group students who were taught by teachers who did not participate in the project. Math knowledge was assessed using the California Standards Test

(CST). The evaluator reported that in year 1, after adjusting for baseline differences, no significant difference was found between the treatment group students and the comparison group. In year 2, students of CLIME teachers significantly outperformed the comparison group students. In year 3, no significant differences were reported. The sample sizes for these analyses ranged from 1,290 to 1,444 for the treatment group and from 1,970 to 2,178 for the comparison group. Similar analyses were also run on the grade-level data individually for grades 3–6 in all three years of the project. All of these analytic contrasts, except for the analysis of students in grades 5 and 6 in year 1, met our rigorous criteria. Among the grade-level comparisons, the evaluators reported that in year 1, the project’s 4th graders performed significantly lower on the test than the comparison group students. In year 2, the treatment group students performed significantly better than the comparison group students in grades 4, 5, and 6. In year 3, no significant differences were found between the two groups at any grade level.

Stanislaus County Math Partnership

State (APR ID): California (CA100182)

Partners: 9 California school districts, Stanislaus County Office of Education, and California State University, Stanislaus

Project Director: Jan Wood

Number of Participants: 92 5th or 6th grade mathematics teachers

Background: The goals of the Stanislaus County Math Partnership were to improve the math proficiency of students in grades 5 and 6 by deepening teachers' content knowledge and instructional strategies through improved professional development and to enhance students' confidence in and attitude towards mathematics.

Description of Professional Development: During 40 hours of intensive summer training, teachers were taught mathematical content knowledge by college faculty. Instructional coaches enhanced teachers' pedagogical knowledge and trained them in how to develop instructions in response to student assessment data. The teachers also participated in 32 hours of lesson study focused on algebra or rational numbers and receive 12 hours of individual instructional coaching.

I am going (back) to school w/ a full bag of tools to use.

I feel more confident in my ability to show multiple representations.

—Participating teachers

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The scope of this evaluation was expansive, covering multiple grade levels, years of the program, and assessment types. Because of the large number of analyses conducted in this report, we were not able to review *all* of the report's findings. Below for each outcome, we present one set of findings that met the criteria.

Teacher Content Knowledge

The evaluators assessed differences between treatment and comparison teachers' content knowledge in all three years of the program, using total scores and subtests from the Diagnostic Teacher Assessments in Mathematics and Science (DTAMS) Assessment. One analysis that met the MSP criteria for rigor assessed whether 105 5th and 6th grade teachers who participated in the program performed better on the conceptual understanding portion of the DTAMS rational numbers subtest than 54 comparison group teachers, who performed similarly to participant teachers on a pre-test. The evaluator reported no significant difference between the two groups in the first year of the program at post-test.

Student Achievement

The evaluators examined differences between students taught by SCMP teachers and those taught by comparison group teachers over the three years of the program, using total score and cluster scores from the California Standards Test (CST) and Ceres Unified School District SCMP Mid-Year assessment. One analysis that met the MSP criteria assessed whether 5th grade students taught by SCMP teachers outperformed comparison students with similar levels of math knowledge at pre-test who were taught by teachers who did not participate in the project. The evaluator reported that while all of the 5th graders had lower scores at the year 1 post-test than at pre-test, the 904 treatment students showed a larger decrease in the total post-test score compared to the 510 comparison group students.

Downey Opportunities in Mathematics Partnership (DO Math)

State (APR ID): California (CA100183)

Partners: 1 school district and the University of California, Los Angeles

Project Director: Melissa Canham

Number of Participants: 45 grade Algebra I teachers

Background: The DO Math project aimed to increase the percentage of students scoring at proficient or advanced levels on the Mathematics test in 7th grade and increase the percentage of English learners (ELs) and students with disabilities taking the Algebra 1 CST in 8th grade by improving the mathematical competence and confidence of the teacher workforce, and upgrading the quality of university mathematics programs for elementary and middle school teachers

Description of Professional Development: Intensive 60-hour summer workshops were provided to teachers in grade 3 Algebra I that focused on number sense in the first year, algebraic thinking in the second year, and geometric thinking in the third year. Follow-up activities were provided during the academic year totaling at least 24 hours, in which teachers applied the content addressed in summer courses and participated in professional learning communities. The follow-ups were broken up into two lesson study cycles. The first cycle included an all-day training facilitated by both the math coach and a UCLA instructor in which teachers reviewed material and pedagogy learned during the summer institutes, looked at district math growth assessment data, and helped groups plan their lessons for their lesson for study. On a second day groups went out and taught, observed, and reflected upon the lesson with the math coach. The same process was repeated during the second cycle. The only differences were that teachers collected pre-test data on their students to help with the planning of the second lesson.

Description of Evaluation with Rigorous Design: The evaluation of *student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The finding that met the criteria is described below.

Student Achievement

The evaluator assessed whether EL students in grades 3–11 who were taught by treatment teachers outperformed comparison group students who were taught by teachers who did not participate in the project. Math knowledge was assessed using the California Standards Test (CST). The evaluator reported no significant difference among the EL students of the two groups on the CST percentile rank in 2012–2013.

Addressing Learning Gaps in Education by Raising Algebraic Understanding and Increasing Content knowledge (ALGEBRAIC) Project

State (APR ID): California (CA100188)

Partners: Lake Elsinore Unified School District, Temecula Valley Unified School District, and San Diego State University's (SDSU) Education and Math Departments

Project Director: Kelli Wise

Number of Participants: 80 teachers in grades 3 through Algebra from 29 schools, primarily high need

Background: The ALGEBRAIC project's aimed to increase teachers' content and pedagogical content knowledge by institutionalizing effective mathematics professional development practices and building professional learning communities. The ALGEBRAIC project seeks to increase student understanding and achievement to help prepare them for challenging mathematics courses.

Description of Professional Development: In each of the three years, SDSU provided teachers with five 8-hour days of intensive summer training, which centered on algebra topics such as rational numbers, ratio and proportions, algebra and functions, and mathematical reasoning. Teachers actively learned mathematics individually, from students, and from each other. Mini-institutes were conducted twice during each academic year by the SDSU faculty and math coaches/facilitators, for an additional 20 intensive hours. During these mini-institutes, teachers worked in grade level teams, choosing and focusing on grade-specific state math content standards, and developed lesson studies, building on the content and pedagogical strategies from the summer institutes. Presenters also modeled strategies for teaching mathematics to English learners. Video recordings and artifacts from the professional development provided additional data to teachers, and will also serve as a learning tool for SDSU students in their courses.

Conversations about student learning have motivated me to consistently teach lessons that are designed around student inquiry, problem solving and evaluation.

Prior to participation in lesson study, I rarely asked students to pair/share. Now it is part of our daily practice.

—Participating teachers

Description of Evaluation with Rigorous Design: The evaluation of student achievement, using a quasi-experimental design, met the rigorous criteria that determine whether an evaluation was conducted successfully. The scope of this evaluation was expansive covering multiple grade levels, years of the program, and subgroups of students. Because of the large number of analyses conducted in this report, we were not able to review all of the report's findings. Below, we describe the scope of the evaluation and one set of findings that met the criteria.

Student Achievement

The evaluator conducted several analyses to assess whether increases in mathematics knowledge were greater for the treatment students in grades 5 through 12 who were taught by teachers that participated in the project than for a comparison group of students with similar levels of mathematics knowledge at pre-test who were taught by teacher that did not participate in the project. Mathematics knowledge was assessed using the Identifying Needs: Standards Proficiency Exams for California Teachers (INSPECT) and the California Standards Test (CST). One analysis that met the MSP criteria for rigor found that, after adjusting for baseline differences, 463 5th grade treatment students significantly outperformed 273 comparison group students on the grade-specific INSPECT mathematics assessment at post-test.

Inventing, Designing, Engineering, Activities in Science (IDEAS) Project

State (APR ID): California (CA100190)

Partners: Cajon Valley Unions School District, Lakeside Union School District, Santee School District, San Diego County Office of Education, San Diego State University, WestEd/K–12 Alliance

Project Director: David Tupper

Number of Participants: 70 3rd–8th grade science teachers

Background: The IDEAS project aimed to increase teachers’ science content and pedagogical content knowledge and strategies, ultimately to increase student conceptual understanding and achievement in science. Project partners worked together to build a professional learning community, emphasizing technology integration and inclusion of diverse learners.

Description of Professional Development: Teachers participated in 84 hours of professional development including a summer intensive workshop and teaching learning collaboratives (TLC) and after-school sessions during the academic year. All teachers participated in a 5.5 day (44 hours) summer workshop, which focused on science content, pedagogy, and leadership strategies. University faculty, districts teachers, and other partners delivered the professional development, modeling pedagogical strategies such as student-to-student discourse and project based learning. During the academic year, participants engaged in two rounds of team-based grade-specific TLC series, a specialized form of lesson study, in which teams of 4 teacher participants co-design, co-teach, and debrief a lesson by analyzing student data, then redesign, co-teach the lesson to another set of students, and finally reflect on instructional and assessment practices that worked particularly well. Teachers also participated in after-school sessions in which they analyzed student data and design classroom interventions. Finally, six district teachers received training to become teacher leaders and work with the lesson study teams.

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The finding that met the criteria is described below.

Teacher Content Knowledge

The evaluators assessed whether increases in science knowledge were greater for the 46 teachers in grades 3 through 8 who participated in the project than for the 31 comparison teachers with similar levels of science knowledge at pre-test who did not participate in the project. Science knowledge was assessed using a state-required science test. The evaluators reported that the post-test scores of teachers participating in the MSP project were significantly higher than those of comparison teachers.

Inquiry Core for 3rd, 4th and 5th

State (APR ID): California (CA100192)

Partners: Paso Robles Joint Unified School District, California Polytechnic State University, and eight additional school districts and one private school.

Project Director: Kelly Roth

Number of Participants: 61 3rd–5th grade science teachers

Background: The goal of the “Inquiry Core for 3rd, 4th and 5th” partnership was to provide teachers intensive content knowledge and additional instructional strategies for optimal student understanding and increased student achievement. Participating teachers and their students in this program used inquiry based instruction strategies, hands-on activities promoting language development and higher thinking-skills strategies. Standard-based benchmark assessments of key student science concepts were expected to be developed and administered over the three-year project.

Description of Professional Development: California Polytechnic State University provided and supported three grade-specific summer sessions, totaling 60 hours, to increase teacher content knowledge and student conceptual understanding and achievement. These sessions modeled inquiry-based hands-on science instruction, links to English language development, and classroom planning. Each year centered on a different content-topic standard. In its third year, the project focused on life science content standards. Participants engaged in 24 hours of academic-year follow-up activities. Lesson study was used to provide teachers the opportunity to work on career advancement through collaboration and researching student learning.

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluators assessed whether 16 5th grade teachers who participated in the program showed greater increases in content knowledge than a comparison group of 24 teachers who did not participate in the program. Teacher content knowledge was assessed using surveys composed of items from the Study of Instructional Improvement. The evaluators reported that the program had a positive impact on MSP teachers’ content knowledge after the training, with post-test scores significantly higher for treatment teachers than for comparison group teachers.

Project DELTA (Developing Educators Learning to Teach Algebraically)

State (APR ID): California (CA100193)

Partners: 5 school districts, the Riverside County Office of Education (RCOE), California State University San Bernardino, and Key Data Systems.

Project Director: Gwen Hancock

Number of Participants: 73 3rd – 8th grade mathematics teachers

Background: In its third year, Project DELTA was focused on functions and equations, with the goals of increasing teachers' math content knowledge and pedagogy and increasing the number of highly subject-qualified teachers. On the student side, the project sought to increase student advanced math course enrollment, achievement against matched comparisons, and disposition and attitude toward mathematics, as well as to decrease Algebra I repeaters.

Description of Professional Development: Provided by California State University San Bernardino math faculty, RCOE math coordinators, and a Common Core expert, the professional development consisted of an 8-day 64-hour summer institute and 48 hours of follow-up. The professional development had a general focus on functions and equations, but the choice of topics was guided by the new Common Core math standards. To develop teachers' mathematical thinking and classroom instruction, the providers used concrete objects, such as ropes to develop concepts of linear functions or balloons to model volume functions. Two Saturday sessions consisted of topics chosen based on needs assessments, including developing Common Core lessons with rigor and increasing cognitive demand. Additionally, follow-up hours centered on the lesson study model.

Description of Evaluation with Rigorous Design: The evaluation of *student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Student Achievement

The evaluators assessed whether increases in math knowledge were greater among treatment students in grades 3–8 who were taught by DELTA teachers than among a group of comparison students with similar levels of math knowledge at pre-test who were taught by teachers who did not participate in the program in 2012 and 2013. The number of students included in the analyses ranged from 1,961 to 2,267 for the treatment group and from 2,010 to 2,065 for the comparison group. Differences in math knowledge were measured using the state assessment test, the California Standards Test (CST). The evaluators reported that there was no difference in the percent proficiency between the DELTA students and the comparison group students on the post-test in either year.

Additionally, the evaluators assessed whether increases in math knowledge were greater among 506 students who were taught by DELTA teachers who took extra coursework for certification (MCPT) than among a group of 3,771 comparison students in 2012. The evaluators reported that there was a significantly greater increase in the percent proficiency of the students in MCPT classrooms than comparison group students.

Algebra I Model Curriculum Field Study

State (APR ID): Connecticut (CT120701)

Partners: 25 Connecticut school districts, six additional districts in partnership with Education Connection, and Central Connecticut State University

Project Director: S. Louise Gould

Number of Participants: 55 Algebra I curriculum teachers from 52 schools that represent a range of urban, suburban, small town, and rural schools

Background: In 2008, the Connecticut State Department of Education (CSDE) worked with a variety of partners to develop a plan for secondary school reform. The plan focused on increasing student engagement, 21st Century skills, and rigorous content, and called for state-mandated curricula for Algebra I, Geometry, Algebra II, and Probability/Statistics. Algebra I was the first of these curricula to be developed, and was written by a team of Connecticut high school teachers and college professors in 2009. The Math and Science Partnership funded a three-year pilot study of the Algebra I curriculum. The project implemented a professional development plan to support three cohorts of algebra teachers. The curriculum was revised in response to teacher feedback and was aligned with Connecticut's new CCSS.

Description of Professional Development: Teachers attended one of two summer institutes offered, each of which had eight units. There were three presenters, at least one of whom was a higher education faculty member, and one of whom was an experienced classroom teacher. Participants also attended February and June Saturday morning sessions. Small groups of teachers could voluntarily meet for a virtual discussion group (Google hangouts) lead by a teacher mentor. Participants attended an average of 29 professional development hours, but there was considerable variability in the number of hours attended. Participants were provided with professional development kits that included materials such as lesson and implementation plans, guidelines for meeting the needs of English language learners, and tips for managing groups in a classroom.

Description of Evaluation with Rigorous Design: The evaluation of *student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Student Achievement

The evaluators assessed whether increases in math knowledge were greater for 1,686 middle and high school students who were taught by teachers in the project than for an unmatched comparison group of 1,542 middle and high school students with similar levels of math knowledge at pre-test who were taught by teachers that did not participate in the program. Math knowledge was measured using an assessment that was developed by experts for the project and was aligned to the program curriculum, drawing mostly from items on state and national assessment tests. The evaluators reported that, after adjusting for baseline differences, the students in CCSS Algebra I teachers' classroom, on average, scored higher than comparison group students on the year-end math test.

Using the same assessment, evaluators also examined whether increases in math knowledge were greater for 122 students who were taught by Algebra I teachers participating in the Algebra21 component of CCSS Algebra I curriculum than students in two comparison groups (122 matched students and 1,542 unmatched students) who were taught by non-participating teachers. The evaluators reported that, after adjusting for baseline differences, the students in Algebra21 teachers' classrooms did as well as students in the matched and unmatched comparison group on the year-end math test.

Math and Science Endorsement Cohorts

State (APR ID): Georgia (GA110629)

Partners: Gwinnett County Public Schools and the Georgia Institute of Technology STEM faculty and graduate students

Project Director: Jonathon Wetherington

Number of Participants: 19 elementary science teachers and 11 elementary mathematics teachers

Background: The goal of this project was to increase the number of elementary teachers with the Georgia Professional Standards Commission approved mathematics and/or science endorsement, and to increase the content knowledge of elementary math and science teachers more generally. Participants in the science endorsement cohort took college-level courses on science content and pedagogical strategies for science instruction, while participants in the math endorsement cohort took courses on math content and pedagogical mathematics for early childhood education. The project also sought to increase student achievement in 3rd–5th grade mathematics and science, as well as in the English language learner and students with disabilities subgroups.

Description of Professional Development: Delivered by Master Teachers and higher education faculty, the professional development was a course-based model, encompassing four content-focused courses that each lasted 50 hours. Through the sequence of courses, teachers had opportunities for on-site training, field work, on-line discussions and learning opportunities, and classroom observations. Activities and lessons were chosen based upon content areas of weakness. Teachers were provided with content texts and lab supplies to enhance the connection between content knowledge and instruction. Upon course completion, teachers delivered two lesson plans demonstrating mastery of content and pedagogy.

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The finding that met the criteria is described below.

Teacher Content Knowledge

The evaluators assessed whether increases in science knowledge were greater for the 19 elementary school teachers who participated in the program than for the 21 comparison teachers with similar levels of science knowledge at pre-test who did not participate in the program. Science knowledge was assessed using the Misconception Oriented Standards-based Assessment in Science (MOSART). The evaluators reported that the post-test scores of teachers participating in the program were not significantly higher than those of comparison teachers, after adjusting for pre-test scores.

Graduate / Bradley University's Environmental Science Education Partnership Program

State (APR ID): Illinois (IL120903)

Partners: Peoria Public School District 150, Bradley University's STEM and Education faculty

Project Director: Kelly McConnaughay

Number of Participants: 15 middle and high school teachers

Background: By providing teachers with graduate training in science, this research-based professional development project aimed to enrich teachers' content knowledge and pedagogical expertise, particularly in environmental science. The program had three primary goals: 1) to increase teacher content and process knowledge in environmental science; 2) to increase teachers' ability to use inquiry-based curricula and instructional techniques and formative assessment and action research techniques; and 3) to develop teacher leadership skills.

Description of Professional Development: Participants completed 33 hours of graduate courses, which were taught by Bradley University's STEM or education faculty and were guided by research findings on knowledge growth and retention. Courses were inquiry-based, providing participants with hands-on research and reporting experience, including a 10-week summer opportunity to work with a practicing scientist and a final-year action research project to evaluate an inquiry-based lesson for use in their classrooms or for professional development of other teachers. Participants were given personal profiles tracking their progress so that they could monitor and assess their own growth.

Description of Evaluation with Rigorous Design:

The evaluation of *teacher content knowledge*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

Evaluators assessed whether the 15 teachers participating in the project and 11 comparison group teachers, who did not participate in the project, showed increases in math content knowledge between the five time points the teachers were tested. Math content knowledge was assessed using a test developed by the Bradley faculty. The evaluators reported that the project teachers had a significant increase in their content knowledge in four out of the ten comparisons made and the comparison group teachers had a significant increase in five out of the ten comparisons made.

WIP 3 ROE 45 MRI Math Connections II

State (APR ID): Illinois (IL121016)

Partners: Monroe Randolph Regional Office of Education, Southern Illinois University-Edwardsville, Chester Community Unit School District 139, and the Institute for Mathematics and Education (IM&E) at the University of Arizona

Project Director: Mary Ann Quivey

Number of Participants: 51 elementary, middle school, special education and appropriate secondary teachers in southern Illinois

Background: Monroe Randolph Intel Math Connections worked with southern Illinois teachers. The project primarily served to improve math content expertise and pedagogy for elementary teachers, though special education and secondary teachers were involved as well. One of the project goals was to enhance the mathematical content knowledge (especially in number sense, algebra, and geometry), problem-solving ability, assessment techniques, and instructional abilities of these teachers. The project also aimed to build a professional learning community for southern Illinois teachers, provide these teachers with resources to enhance inquiry learning and critical thinking in their classrooms, encourage teachers to conduct action research, and develop the leadership skills of these teachers.

I think we have all seen firsthand the difference between learning a process and understanding a concept. I have gained an appreciation for the connections between number sense, algebra and geometry which start in Kindergarten. I also have to say the instructors were all very good at practicing the instructional strategies involved with the common core standards. They modeled things I will now use in my classroom. I see the importance of group work and multiple solutions in teaching concepts.

—2012 cohort teacher

Description of Professional Development:

Mathematicians, math educators, and math coaches taught eighty hours of professional development, divided into eight units, based on the Intel Math curriculum. Participants attended a summer institute and four follow-up days of instruction, and were offered district-based team meetings. Instruction consisted of concept development, individual practice, small group discussion, whole group instruction, and homework. Formative assessment techniques were used extensively and teachers were expected to find multiple problem-solving solutions.

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge, teacher classroom practice, and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluator assessed whether 31 elementary school teachers participating in the project showed greater increases in math content knowledge than a matched comparison group of 31 elementary school teachers, who did not participate in the project. Teacher content knowledge was assessed using the Intel Math Content assessment. The evaluator reported that, after adjusting for baseline differences, the project teachers had a significantly greater increase in their content knowledge than the comparison group teachers.

Teacher Classroom Practice

The evaluator also examined whether 31 elementary school project teachers improved their classroom practice more than a matched comparison group of 31 elementary school teachers who did not participate in the project. Classroom practice was assessed using a math curriculum survey which was

derived from the Surveys of Enacted Curriculum—Mathematics. The evaluator reported that, after adjusting for baseline differences, the elementary project teachers did not have significantly larger increases in their reasoning and problem solving, educational technology, and collegiality composite measures (examining the extent to which mathematics teachers in the school regularly observe each other teaching classes) than the comparison group teachers.

Student Achievement

The evaluator assessed whether increases in math knowledge were greater among 24 4th graders who were taught by teachers participating in the project than among a group of 37 comparison students with similar levels of math knowledge at pre-test, who were taught by non-participating teachers. Math knowledge was assessed using the ThinkLink testing system. The evaluator reported that, after adjusting for baseline differences, the project's 4th graders did not show any increases on the tests while the comparison group had significant increases in their math knowledge. There was a significant difference between the comparison group and the project group on the post-test but the evaluator concludes that there was no evidence to support that the difference was due to the project.

Promoting Teacher Quality and Student Achievement in Science

State (APR ID): North Carolina (NC100612)

Partners: Catawba College, Rowan-Salisbury Public School System

Project Director: Lisa Wear

Number of Participants: 127 teachers from 26 public elementary and middle schools, and 3 nearby private schools

Background: The Promoting Teacher Quality and Student Achievement in Science project provided comprehensive, content driven, and standards based professional development in physical science for teachers in grades K–8 to improve their content knowledge, pedagogical instruction, use of inquiry learning, integration of technology in instruction, and alignment with Science Essential Standards to ultimately enhance their students’ science mastery. The project tried to foster sustained relations between teachers and disciplinary faculty.

Description of Professional Development: During the two-week summer institute, four college STEM faculty and science specialists taught lessons about force and motion, matter, chemical and physical changes, energy, conservation, and transfer. These sessions were attended by teachers and principals and used inquiry-based learning techniques. Teachers attended four Saturday follow-up sessions throughout the academic year and participated in science leadership teams.

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluators assessed whether 67 teachers in grades K–8 in the second cohort of the program showed greater physical sciences content knowledge than a comparison group of 61 teachers who did not participate in the program. Teacher content knowledge was assessed in summer 2012 using pre-test and post-test scores on a novel instrument created for the evaluation using items from the Misconception Oriented Standards-based Assessment Resource for Teachers (MOSART) Physical Science tests, Trends in International Mathematics and Science Study (TIMSS) Science Test, and state End-of-Course or End-of-Grade science tests. The evaluator compared the gain scores for the treatment and comparison teachers and reported that the gain in content knowledge from pre-test to post-test was significantly higher for program teachers than the comparison teachers.

Assessing Core Content and Ensuring Success in Science (ACCESS)

State (APR ID): North Carolina (NC100613)

Partners: Catawba County Schools, North Carolina State University's College of Physical and Mathematical Sciences and the Science House and Appalachian State University's North Carolina Center of Engineering and Technology

Project Director: Carol Moore

Number of Participants: 150 K–5 teachers from all 16 elementary schools in the district

Background: A research-based approach to professional development, ACCESS aimed to enhance teacher content and pedagogical knowledge in science and to build school-based leadership and support for the science program. Additionally, through training teacher leaders and through sustainable partnerships, the project hoped to establish a continued source of professional development to improve science instruction and student understanding of key concepts.

Description of Professional Development: With regular involvement of university faculty in planning and implementation, 30 participants received 305 hours of instruction and support to become teacher leaders, and 120 participants received at least 90 hours of instruction and support. Twenty-eight teachers participated in a two-week summer institute that concentrated on core physical science content, hands-on inquiry-based pedagogy, conceptual change, and technology integration. Two full-day follow-up sessions covered additional topics such as science content, safety, and development of professional learning communities (PLCs), which consisted of a teacher leader and at least four teachers in their schools. To strengthen administrative leadership and support, the project held workshops for school administrators to help them support their school's science program.

My approach has changed to allowing the students to investigate on their own & I have learned to ask better questions. I am doing more science journals, investigations, and free exploration. I no longer give students info to start. I give them materials and have them explore and discuss. I have become more relaxed about the noise and movement in the room. Students are free to roam and discuss. This allows conversation to begin and new ideas are shared. The investigation stimulates thinking to wonder "why" and "what if" which continued with deeper thinking. This approach engages all kids, & no one feels they need to be "right." All students feel successful.

—ACCESS Participant

Description of Evaluation with Rigorous Design:

The evaluation of *teacher classroom practices*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Classroom Practice

The evaluator tested whether the 29 elementary teachers in the teacher leader group, who received a series of professional development components, differed from a comparison group of 46 teachers in the PLC group, who only received one professional development component, the PLC component of the intervention, in the frequency of using certain classroom practices. Specifically, the frequency of using traditional teaching strategies, standards-based teaching strategies, and formative and summative assessment strategies was assessed using measures developed specifically for the study. The evaluator compared gain scores for the treatment and comparison teachers and reported that gains in using standards-based strategies and using formative and summative assessments were greater for the teacher leader group than for the comparison group. The evaluator also reported that the frequency of using traditional teaching strategies declined more for the teacher leader group than for the comparison group, and the difference nearly reached statistical significance.

Curriculum Topic Study to Enhance Achievement in Mathematics and Science (C-TEAMS)

State (APR ID): New Jersey (NJ100717)

Partners: 16 school districts and nonpublic schools, Stevens Institute of Technology, and St. Peter's College

Project Director: Philip Leopold

Number of Participants: 30 math and 36 science teachers in elementary through high school

Background: The goal of the curriculum topic study was to deepen multi-grade groups of teachers' science and math content knowledge, pedagogical content knowledge, and understanding of New Jersey state standards through professional development and a professional learning community. Ultimately, it sought to increase student achievement in science and math in all grades.

Description of Professional Development: Teachers participated in 116 hours of professional development including a 70-hour summer institute, three workshops during the academic year (two in-person and one on-line), a professional learning community, and a teacher exchange program. They also received in-class coaching.

Teachers reengineered their math and science curriculum during the course of this professional development. The math and science content topics covered included earth science, statistics, probability, geometry, and computer-based technology.

Description of Evaluation with Rigorous Design: The evaluation of *teacher content knowledge and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

Evaluators assessed whether the 24 C-TEAMS teachers in grades K–12 and 12 comparison group teachers, who did not receive the C-TEAMS professional development, showed pre-to-post increases in math and science content knowledge in the 2011–2012 academic year. Content knowledge was measured with the Diagnostic Teacher Assessments in Mathematics and Science (DTAMS) Algebra and Physical Science subtests. The evaluators reported that neither the C-TEAMS teachers nor the comparison group teachers had a significant increase in their content knowledge between pre-test and post-test.

Student Achievement

Evaluators assessed whether increases in math and science knowledge in the 2012–2013 academic year were greater among middle school students who were taught by C-TEAMS teachers than among a group of comparison students with similar levels of knowledge at pre-test who were taught by teachers that did not participate in C-TEAMS. Math knowledge was assessed using the Michigan Mathematics Leadership Academy assessment, and science knowledge was assessed using tests constructed from released items from national and international standardized assessments, state standardized assessments, and the Misconception Oriented Standards-Based Assessment Resources for Teachers (MOSART). The number of students included in the analyses ranged from 32 to 920 for treatment students and from 21 to 970 for comparison students. The evaluators reported that the overall pre-to-post change in math scores for treatment students was significant, while the students in

As a classroom teacher I benefited from engaging in CTS because it showed me an easy way to group related topics of science content which would then make it easier for me to help my students make those connections. Teaching content areas in isolation from each other is the wrong approach. CTS enable students to make the connections and build upon prior knowledge to gain a wider understanding of the world they live in.

—C-TEAMS Science teacher

the comparison group had no statistically significant increase on the overall score. However, when the analysis was done by grade level, both 6th and 7th grade students in the comparison groups showed statistically significant increases in their math scores. For the science test, the overall pre-to-post changes in test scores for grades 6, 7, and 8 for treatment students and students in the comparison group were all significant.

Greater Allentown Math Science Partnership

State (APR ID): Pennsylvania (PA100728)

Partners: Allentown School District, 4 regional colleges, and the Da Vinci Science Center

Project Director: Kelly Murray

Number of Participants: 90 teachers in grades 1–8 in Allentown School District were randomly assigned to the math, science, or control condition

Background: With proficiency rates below 75 percent in many schools involved in this project, the Greater Allentown Math Science Partnership aimed to address Allentown School District’s significant need to improve student achievement in math and science. It sought to increase middle school teachers’ understanding of math and science content and their use of research-based instructional practices.

Description of Professional Development: University faculty members in math and science content and in education, K–12 educators, and other experienced professional development personnel led the professional development program. This included an 80-hour August workshop in math or science, 16–32 hours of academic-year professional learning community (PLC) activities, 8 hours of on-line PLC work, and optional individualized classroom and curriculum coaching by project staff. The summer workshops modelled inquiry learning, aligned with Allentown’s curriculum and with the standards. Topics covered included physical science, life science, and earth science, algebra, geometry, probability, and statistics.

Description of Evaluation with Rigorous Design: The randomized control trial examining effects on *teacher content knowledge, teacher classroom practice, and student achievement* met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluator assessed whether treatment teachers and control group teachers, who did not participate in the program, showed significant increases in math and science content knowledge. Content knowledge was assessed using the Diagnostic Teacher Assessments in Mathematics and Science (DTAMS). The evaluation examined changes in DTAMS scores between treatment and control teachers in the areas of probability and statistics, geometry, life sciences, and physical sciences, with the number of teachers varying by topic. The number of treatment teachers per topic ranged between 15 and 24, and the number of control teachers per topic ranged between 11 and 23. The evaluator reported that the math treatment teachers had significant increases in their math content knowledge as a result of the two-week intensive program in the last two years of the program. The science treatment group had a significant increase in their content knowledge in the second year, with no significant differences in the first year.

Teacher Classroom Practice

The evaluator assessed whether the treatment teachers in grades 1–8 improved their classroom practice more than the control group teachers using the Reformed Teaching Observation Protocol (RTOP). The evaluator reported that there was no significant difference between the two treatment groups (24 math teachers and 27 science teachers) and the control group (22 teachers) in the final year.

Student Achievement

The evaluator assessed whether increases in math and science knowledge were greater among students in grades 3–8 who were taught by treatment teachers than among students taught by control group teachers. Evaluators examined differences over three years of the program from 2010–2011

through 2012–2013, with the number of students varying by year. The number of treatment students per year ranged between 47 and 892, and the number of control student per year ranged between 228 and 880. Content knowledge was assessed using the state achievement test, the Pennsylvania System of School Assessment. The evaluator found no significant differences between groups in math or science assessments in any of the three years.

School District of Philadelphia Math and Science Partnership

State (APR ID): Pennsylvania (PA100730)

Partners: St. Joseph's University, School District of Philadelphia

Project Directors: Rachel Cherry

Number of Participants: 20 high school teachers

Background: The School District of Philadelphia MSP project sought to better meet the educational needs of its students by advancing teachers' content knowledge and pedagogical practice.

Description of Professional Development: Participants took two 40-hour graduate-level courses over the summer, covering the history of mathematics and mathematical problem solving. Additional courses on effective practices in teaching mathematics and advanced geometry content and teaching methods were also offered. During the academic year Saturday workshops and fellowship meetings were offered to increase science teachers' understanding of physical science and math teachers' grasp of geometry, linear algebra, and mathematical thinking skills. As an additional benefit, math teachers were offered the opportunity to join the National Council of Mathematics Teachers using money from the grant and were offered the opportunity to take additional teacher effectiveness classes at St. Joseph's University.

Description of Evaluation with Rigorous Design:

The evaluation of *teacher content knowledge*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluators assessed whether a group of math teachers in grades K–12 who attended the program showed greater increases in content knowledge than a comparison group of teachers who did not participate in the program. Teachers' algebra knowledge was assessed using the Misconception Oriented Standards-Based Assessment Resource for Teachers (MOSART) in the first year of the program and a project-developed assessment on mathematical problem solving and the history of mathematics in the third year of the program. The number of teachers included in these analyses ranged from 9 to 20 treatment teachers and from 14 to 21 comparison group teachers. The evaluators reported that in the first and third years, the program had a positive impact on teachers' disciplinary content knowledge.

Project ARRMS: Achieving Rigor and Relevance in Math and Science

State (APR ID): Pennsylvania (PA100731)

Partners: Lancaster-Lebanon Intermediate Unit 13, a local college and a local university, and industry partners

Project Director: Joey Rider-Bertrand

Number of Participants: 53 secondary math and science teachers in 19 public schools where 25 percent or more of students scored “basic” or “below basic” on statewide assessments, 1 charter school, and 7 nearby nonpublic schools

Background: Paralleling trends in the country at large and in Pennsylvania, Lancaster and Lebanon counties in Pennsylvania struggled with maintaining middle and high school students’ proficiency in math and science. Project ARRMS aimed to develop a professional development framework wherein secondary math and science teachers could deepen their content knowledge, improve the rigor and relevance of the lessons they teach, network with local and regional leaders in STEM fields, and share best practices across district boundaries.

Description of Professional Development: Teachers participated in an 80-hour (two-week) summer institute to enrich their content knowledge and 24 follow-up hours of individualized and small group instructional coaching with an assigned instructional coach. The summer institute was led by university faculty. Teachers had various STEM field and laboratory activities and were trained to use laboratory equipment. Participants also collaborated and had virtual discussions in an on-line community.

Description of Evaluation with Rigorous Design:

The evaluation examining effects on *teacher content knowledge, teacher classroom practice, and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

Evaluators assessed whether in the second year of the program, 28 treatment teachers, who taught grades 6–12, showed greater increase in science content knowledge than a group of 17 comparison teachers who did not participate in the project. Science content knowledge was assessed using a project-developed test which drew from released and practice items from standardized assessments in science. The evaluators reported that, after adjusting for baseline differences, the difference between the treatment and comparison teachers on the post-test and follow-up assessments in the second year of the program were not statistically significant but the difference in means was in the expected direction. That is, the means for the treatment teachers were higher than the means for the comparison teachers.

Teacher Classroom Practice

The classroom practices of treatment teachers, who taught grades 6–12, were compared to those of the teachers in the comparison group. Classroom practice was assessed using the Surveys of Enacted Curriculum (SEC) and the Reform Teaching Observation Protocol (RTOP). The number of teachers included in these analyses ranged from 53 to 108 for the treatment teachers and from 27 to 46 for the comparison group teachers.

The evaluators reported that, after adjusting for baseline differences in SEC scores, across the three years, the math treatment teachers’ estimate of the time their students spent demonstrating understanding of mathematical ideas was greater than comparison teachers’ estimate of their students’

time. Also, both math treatment and comparison teachers spent less classroom time performing procedures and using active learning at post-test than at pre-test.

The evaluators also reported that science treatment teachers spent less classroom time performing procedures than the comparison teachers after adjusting for baseline differences. There was also a weak tendency for both treatment and comparison teachers to indicate greater content readiness at post-test than at pre-test.

To summarize the results on RTOP, the evaluators reported that, after adjusting for baseline differences, the treatment teachers across the three years of Project ARRMS showed significantly greater improvements than the comparison teachers in all aspects of teaching measured by the RTOP.

Student Achievement

Evaluators assessed whether increases in math achievement were greater among 105 students in grade 7 who were taught by treatment teachers than among a group of 186 comparison group students who were taught by teachers who did not participate in the project. The treatment and comparison group students had similar levels of math knowledge at pre-test. Math knowledge was assessed using the state achievement test, the Pennsylvania System of School Assessment (PSSA). The evaluators reported that students of treatment and comparison teachers did not differ significantly with respect to their 2013 Math PSSA scaled scores after controlling for baseline PSSA scaled score.

Building the Foundation to Algebra in Grades 3–8

State (APR ID): Pennsylvania (PA100732)

Partners: Intermediate Unit 1 (IU1), Pennsylvania State Fayette Eberly Campus (PSF), 8 school districts in southwestern Pennsylvania that were underperforming, and one non-public school

Project Director: Nancy Tsupros

Number of Participants: 55 3rd–8th grade teachers

Background: Building the Foundation to Algebra aimed to increase teachers' content knowledge about algebra and skills essential to success in algebra. It also taught participants research-based effective teaching practices. This project was modeled in part on a successful MSP project in Kentucky.

Description of Professional Development: Teachers participated in a summer academy that covered algebra and concepts fundamental to algebra (e.g., numeracy, rational numbers, and geometry) over the course of two weeks (80 hours). During the academic year, participants attended three follow-up days during which suggestions for assessment strategies and appropriate tasks and activities were offered. Teachers received graduate course credit for this work. Courses were taught both by PSF math faculty and IU1 staff, and sessions were held at PSF because the university atmosphere facilitated learning. Teachers were shown STEM connections, taught engineering problems, and learned about engineering careers. Teachers were observed in their classroom during the final year of the project.

Description of Evaluation with Rigorous Design: The evaluation examining effects of *teacher content knowledge and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Teacher Content Knowledge

The evaluators assessed whether 50 treatment teachers who taught math in grades 3–8 showed greater increase in math content knowledge than a comparison group of 30 teachers who did not participate in the project. Math content knowledge was assessed using the Diagnostic Teacher Assessments in Mathematics and Science (DTAMS) - Rational Number Concepts and Operations test. The evaluators reported that while the post-test means for the second year treatment teachers were not significantly different than post-test means of the comparison group teachers, the gains made by the treatment teachers between pre- and post-test were significantly greater than the gains made by the comparison group teachers in year 2.

Student Achievement

The evaluators assessed whether increases in math knowledge were greater among 1,220 students in grades 4–8 who were taught by treatment teachers than among a group of 447 comparison students with similar levels of math knowledge at pre-test who were taught by teachers who did not participate in the project. Math knowledge was assessed using the state achievement test, the Pennsylvania System of School Assessment (PSSA). The evaluators reported that, after adjusting for baseline differences, there was a significant difference in the 2013 math scores; the comparison group students performed significantly higher than the students of treatment teachers.

Laurens County STEM Teacher Development Program

State (APR ID): South Carolina (SC111016)

Partners: Laurens County School District 55, Piedmont Technical College, and Joe Adair Outdoor Education Center

Project Director: Jody Penland

Number of Participants: 60 3rd–12th grade math and science teachers

Background: The STEM Teacher Development Program aimed to improve teachers' content knowledge and self-efficacy, raise student's state test scores, and close achievement gaps. Rigorous content training was tailored to the specific needs of teachers and their students.

Description of Professional Development: Participants took part in 100-hours summer institutes focused on STEM content and instructional strategies and aligned with state standards, as well as ongoing professional development throughout the academic year. Teachers also had the option to take graduate courses on STEM content, and they were provided with tools to use in their current lessons to engage students and teach using hands-on methods.

Description of Evaluation with Rigorous Design:

The evaluation of *student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

Student Achievement

Evaluators compared the math achievement scores of students of treatment and comparison teachers who self-selected into these conditions. This analysis was conducted for students in grades 3–8 from the 2011–2012 and the 2012–2013 academic years. Student achievement was assessed using the state's Palmetto Assessment of State Standards (PASS). The number of students included in these analyses ranged from 794 to 966 for the treatment group and from 971 to 1,034 for the comparison group, depending on the year. The evaluators reported significantly larger gains for students of program teachers than students of comparison teachers in both academic years.

Raising Achievement in Mathematics through Fostering Algebraic Thinking (RAM t-FAT)

State (APR ID): Wisconsin (WI110908)

Partners: West Allis-West Milwaukee School District

Project Director: Marta Magiera

Number of Participants: 30 middle and high school math teachers

Background: The RAM t-FAT project was guided by the ideas that 1) math instruction should be based on interpretation of math more than doing math, 2) some math teachers struggle to teach problem-solving, and 3) collaborative activities can enhance teachers' learning. The project aimed to create a vibrant mathematics teacher learning community characterized by respect and by norms for critical dialogue about mathematics teaching and learning. The primary content focus was on algebra and algebraic thinking.

Description of Professional Development: Each year of the program, site-based professional development activities involved monthly sessions conducted after school hours. Off-site based professional development included an 80-hour summer institutes that was broken up into one week in June and one week in August. Teachers also received mentoring and peer coaching and could participate in an on-line community. The professional development sought to help create a coherent picture of mathematics instruction throughout the district and to support vertical and horizontal collaboration among the participants.

Description of Evaluation with Rigorous Design:

The evaluation of *teacher content knowledge and student achievement*, using a quasi-experimental design, met the rigorous criteria used to determine whether an evaluation was conducted successfully. The findings that met the criteria are described below.

The combination of new resources, discussing theory vs. practice, and guidance from the extremely competent staff from Marquette have been invaluable for me as a math teacher. This experience has helped me to improve my teaching practices significantly and made switching to the common core manageable.

—RAM t-FAT teacher

Teacher Content Knowledge

The evaluator assessed whether 29 middle and high school participant teachers scored higher than 10 comparison group teachers on a measure of math content knowledge in 2013. Math content knowledge was assessed using selected questions from the Learning Mathematics for Teaching (LMT) that addressed knowledge of algebra and patterns. The evaluator reported that the participant teachers had greater gains than the comparison group teachers; however, there were no statistically significant differences between the two groups gain scores.

Student Achievement

The evaluator assessed whether increases in math knowledge were greater among middle and high school students who were taught by participating teachers than among a group of comparison students with similar levels of math knowledge at pre-test who were taught by non-participating teachers. The analysis of students from 2011 and 2013 met the MSP criteria. The sample sizes for these analyses ranged from 2,234 to 2,544 treatment students and from 1,445 to 2,366 comparison group students. Math knowledge was assessed using the Measures of Academic Progress (MAP) exam. The evaluator reported that even though the differences in the average gain scores between students of participating and non-participating teachers were statistically significant for both years, they do not consider these differences to be practically significant as the magnitude of the differences is small relative to the standard deviation in the test scores.

Chapter 6: Summary and Conclusions

The MSP program was created in 2001 to fund collaborative partnerships between high-need school districts and mathematics, science, and engineering departments at institutions of higher education (IHEs). Through these partnerships, the MSP program seeks to provide intensive content-rich professional development to teachers and other school staff, thus improving classroom instruction and ultimately student achievement in mathematics and science. Since the program's inception, it has grown to encompass more projects and serve more participants, who, in turn, have served more students. In Performance Period 2012 (PP12), 488 individual MSP projects were in operation throughout the country. These projects provided professional development to over 40,000 educators who taught over 2.3 million students. In some cases, these educators also trained their fellow teachers, thus influencing an even larger number of teachers and students.

In accordance with the legislation, MSP projects established partnerships between school districts and IHEs as well as with a wide variety of other organizations. Over three thousand faculty members from mathematics, science, engineering, and other departments at IHEs were involved with the MSP projects.

Over half of MSP projects (53 percent) in PP12 conducted summer institutes, a model of professional development designed to provide a period of intensive study of STEM content over a relatively short period of time. Nearly all of the projects that offered summer institutes also conducted follow-up activities, with the aim of enhancing or extending the knowledge gained by participants over the summer. Projects that provided summer institutes with follow-up activities provided participants with a median of 97 hours of professional development. Three percent of projects conducted summer institutes with no follow-up. These projects provided participants with a median of 80 hours of professional development. The remaining 47 percent of MSP projects in PP12 primarily delivered professional development during the academic year, with shorter summer sessions often included. These projects also provided participants with a median of 68 hours of professional development.

All projects are required to administer pre- and post-tests during the year(s) in which their teachers were receiving intensive professional development. The most frequently reported assessments of teacher content knowledge in mathematics were standardized tests (67 percent), followed by locally developed tests (34 percent). The use of locally developed assessments to measure teacher content knowledge in science was more prevalent, with 53 percent of projects using locally developed assessments and only 48 percent using standardized assessments. The main advantage of standardized tests is that they have already been tested for validity and reliability, and thus their results can be compared in a normative context. However, standardized tests are not available in all disciplines and are often not well aligned with the context taught. Thus, many projects developed their own assessments to measure growth in teacher content knowledge of the material taught, although they may not have had strong psychometric properties.

Sixty-three percent of participants who were assessed in mathematics showed significant gains in their content knowledge, and 67 percent who were assessed in science showed significant gains in their content knowledge.

The proportion of students taught by MSP teachers who scored at the proficient level or above in state assessments of science remained strong in PP12, while in mathematics the proportion of students that scored proficient or above dropped. In mathematics, 55 percent of students scored at the proficient

level or above, compared to 65 percent and 64 percent in PP10 and PP11, respectively. In science, the proportion of students scoring at the proficient level or above rose to 69 percent from 67 percent the previous two years.

As they work to determine the impact of their programs, many projects are attempting to implement rigorous evaluation designs. Less than 1 percent of projects reported using experimental designs, and 48 percent of projects reported using quasi-experimental designs with comparison groups. However, upon review of the designs of final-year projects, it was found that many of the projects that reported using quasi-experimental designs in fact used one-group designs comparing outcomes for MSP participants between pre- and post-test.

The Criteria for Classifying Designs of MSP Evaluations were initially developed as part of the Data Quality Initiative through the Institute for Education Sciences at the U.S. Department of Education to identify projects that successfully implemented rigorous evaluation designs. These criteria are slightly revised each year in order to bring them into closer alignment with the What Works Clearinghouse Standards (see Appendix B). The criteria were applied to the final evaluation reports of the 71 projects that completed an experimental or comparison group design and submitted complete data. Twenty-one of these projects met the rigorous criteria, which represents over a seven-fold increase from PP08. These 21 projects varied from one another across the types of program offerings, the content area and grade levels targeted, and the number of professional development hours offered.

Ultimately, the success of the MSP program will be determined by the success of its projects in providing effective professional development to teachers across the nation. The MSP program will continue to study the effectiveness of these efforts in order to develop our understanding of what constitutes high quality, effective professional development.

References

- Bobronnikov, E., Sahni, S.D., Fernandes, M., & Bozzi, L. (2013). *A Guide for Reporting on Rigorous Evaluations for the US Department of Education Mathematics and Science Partnerships (MSP)*. Cambridge, MA: Abt Associates Inc. Retrieved July 18, 2014, from <http://www.ed-msp.net/index.php/resources>
- Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2007). *How and Why Do Teacher Credentials Matter for Student Achievement?* (NBER Working Paper 12828). Cambridge, MA: National Bureau of Economic Research.
- Hanushek, E., & Rivkin, S. (2010). Using Value-Added Measures of Teacher Quality. *American Economic Review*, 100(2), 267–271.
- Hill, Heather C., Rowan, B., & Ball, D.L. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*, 42, 371–406.
- Rivkin, S.G., Hanushek, E.A., & John F. Kain, J.F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2), 417–458.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, *WWC Procedures and Standards Handbook*. What Works Clearinghouse. Retrieved July 18, 2014, from <http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19>

Appendix A: Review of Projects with Rigorous Designs

This appendix presents the results of a review of final-year MSP projects submitted in Performance Period 2012 (PP12) that reported evaluating their programs using an experimental or quasi-experimental design. This review sought to determine the extent to which projects successfully conducted rigorous evaluations to yield findings that could be considered to be reliable and valid. To this end, we assessed how project evaluations, as reported in written project evaluation reports, met the criteria established for MSP projects for rigorous evaluations of interventions. We describe how the review was conducted, the criteria used to assess the rigor of projects' evaluations, the results of the review, and recommendations for improving future MSP project evaluations.

Methodology Used for Review

In order to be included in this annual review, projects must indicate in their Annual Performance Reports (APRs) that they submitted a final evaluation report and that they used a comparison-group design. All projects reporting that they submitted a final evaluation report that used a comparison-group design are considered for further review to determine whether they conducted an evaluation that met the MSP standards for a rigorous evaluation. The primary source of information for the review is the final evaluation report. If additional information is required, the review team consults the information provided in the APR, as well as any supplemental materials. If the information required for assessing a criterion is missing from all available documents, then reviewers contact the project director and request it.

The review process occurs in two stages. In the first stage, reviewers verify that the projects are in their final year and use an appropriate comparison-group design that examines outcomes of teacher content knowledge, classroom practices, or student achievement. Since the purpose of the review was to learn about projects that are conducting rigorous impact evaluations, the review was limited to those projects that reported using an experimental or quasi-experimental design, both of which are considered to be appropriate for testing the impact of a program or intervention. Experimental designs, also known as randomized control trials (RCTs), include designs where units of analysis (i.e., teachers, classrooms, or schools) are randomly assigned to a treatment or comparison group. Evaluations with quasi-experimental designs (QEDs) also include a treatment and comparison group, but the units of analysis are not randomly assigned to the groups.

In the second stage, reviewers review the full evaluation report to determine whether the study includes an evaluation that meets the criteria the MSP Program uses to determine whether an evaluation is considered rigorous.

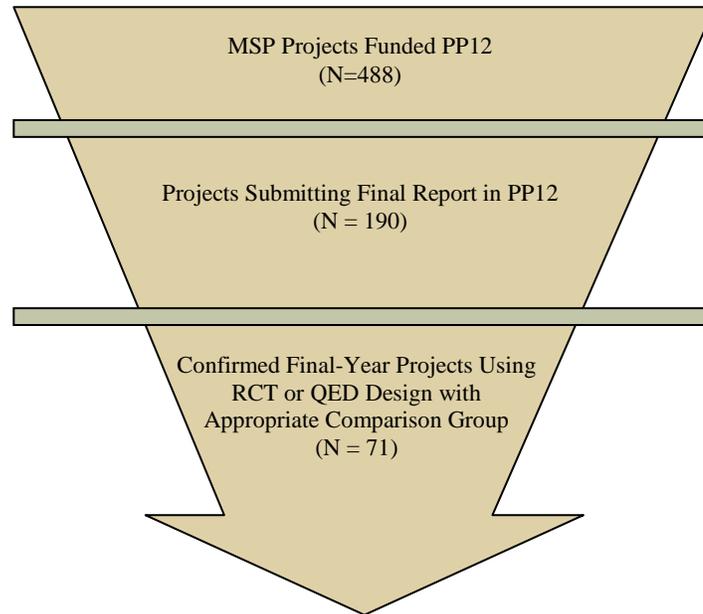
Each of these stages is described in more detail below.

Stage 1 – Defining the Set of Project Evaluations

As seen in Exhibit A.1, among the 488 projects funded in PP12, only 190 projects that reported submitting a final report in PP12 were eligible for review. Focusing only on projects that reported using an experimental or quasi-experimental design narrowed the set of projects for review from 190 to 101. In the first stage of review, we further narrowed the set of projects to 71 by excluding those which could not be verified to be in their final year (5 projects) or did not include an appropriate comparison group (27 projects). For example, some projects evaluated pre- and post-test scores for only a treatment group, or compared treatment group scores to established benchmarks that contained

scores from treatment group students. The remainder of the discussion in this appendix focuses on what we learned from reviewing these 71 projects.²²

Exhibit A.1: Sample of MSP Projects



Sources: Final evaluation reports, annual performance reports, and related documents submitted by MSP projects.

Stage 2 – Assessing MSP Evaluations for Rigor

In the second stage of the review process, documents provided from each of the remaining 71 projects were reviewed more closely to determine whether or not the evaluations met the Criteria for Classifying Designs of MSP Evaluations. These criteria were initially developed by Westat as part of the Data Quality Initiative at the Institute for Education Sciences within the U.S. Department of Education and outline the key elements necessary for implementing a rigorous impact design. The criteria are slightly revised each year in order to bring them into closer alignment with the What Works Clearinghouse Standards (see Appendix B). The four criteria used for assessing the rigor of MSP evaluations are listed below, and more information about each criterion is provided in the following pages:

1. Attrition;
2. Baseline equivalence;
3. Quality of measurement instruments; and
4. Relevant statistics reported.

²² Two of the projects that were not eligible for a full review neither submitted a final-year report nor included a comparison group.

Most evaluation reports include findings on multiple outcomes in various domains. We reviewed only those domains with strong theoretical links to MSP’s goals, which included teacher content knowledge, teacher classroom practices, and student achievement. Projects often conducted research on more than one of these three domains, as well as on multiple subgroups (e.g., third and fourth grade teachers and students), measures (e.g., state assessment and researcher-designed instrument), and time points (e.g., year 2 and year 3). In our review each of these “analytic contrasts,” which is a unique combination of assessment, sample, and time point, is assessed separately. In order for a project to meet the MSP criteria for evaluations using rigorous research methods that yield scientifically valid results, it must have at least one analytic contrast that meets all of the criteria, as described in more detail below.

To meet the criteria, evaluations had to satisfy the requirements of each criterion that was relevant to its design. Of the 71 projects reviewed, 21 projects included at least one analytic contrast that successfully met all of the criteria. Of these 21 projects with analytic contrasts that met the criteria, 20 projects employed quasi-experimental designs, and one used an experimental design. Five projects had evaluations that met the criteria in two domains and three projects had evaluations that met the criteria in all three domains. In total, 14 projects successfully studied their program’s impacts on teacher content knowledge, four projects successfully studied impacts on classroom practices, and 14 projects successfully studied impacts on student achievement. In the review that follows, we present the criteria as well as recommendations for future project evaluations.

Assessing Comparability of Treatment and Comparison Groups

The first two criteria were used to assess the comparability of treatment and comparison groups. A key component of a rigorous impact design is a comparable treatment and control group. The more comparable these groups are, the more likely it is that any observed differences between the groups are attributable to the program studied rather than alternative explanations, confounding factors, or biases.

Attrition

Description. This criterion was assessed for *all experimental evaluations* and for *quasi-experimental designs that did not report baseline equivalence on the final analytic sample*. In order to meet this criterion, key post-test outcomes were measured for at least 70 percent of the original sample (treatment and comparison groups combined) and differential attrition (i.e., difference between treatment group attrition and comparison group attrition) between groups was less than 15 percentage points.

Justification. Randomization is a powerful process because it creates groups that can be considered statistically equivalent. Significant sample attrition can bias the evaluation results, because the participants who drop out of the study may differ from those who remain. When this happens, it is unclear whether differences at the outcome can be attributed to the intervention, or is explained by the difference between the groups that existed prior to the intervention.

For quasi-experimental studies, since units of analysis are not randomly assigned to treatment and comparison groups, evaluators must assess the differences between the groups at baseline in order to demonstrate whether or not the groups are comparable. Groups were considered to be comparable if there were no significant differences in variables related to key outcomes. This comparison should be made of the *analytic sample*, which is defined as the sample of participants for which an outcome, or

related measure, was collected at both pre-test and post-test. Therefore participants for whom there is a baseline measure but no follow-up measure should be excluded from this analysis, as they are not part of the analytic sample. If a study did not provide information to assess baseline equivalence of the analytic sample, but could establish baseline equivalence for the initial sample, then it was subject to the *attrition* criterion.

Screening requirements. To meet this criterion, the experimental evaluation must meet the conditions described below:

1. Present evidence that the overall attrition rate was less than 30 percent. Overall attrition refers to the attrition in the full sample (i.e., the participants in the two groups being compared to one another combined) AND
2. Present evidence that the difference in the attrition rates in the treatment and control groups was 15 percent or less.

When attrition rates were not provided in the evaluation, we calculated attrition rates by subtracting the post-test sample size from the pre-test sample size and dividing by the pre-test sample size. If an evaluation failed to provide this information and met all other criteria, coders contacted the project director for the information required to calculate attrition.

Recommendations.

1. Report the number of units of assignment and units of analysis at the beginning and end of the study.
2. If reporting on subgroups, report sample sizes for all subgroups.
3. Implement a plan for keeping sample participants involved with the study. Some successful evaluations reduced attrition by making follow-up data collection as easy as possible—for example, relying on paper tests rather than on-line surveys (which may be more difficult due to the reliance on respondent initiative and reliable Internet access) or using data from mandatory state tests, virtually guaranteeing follow-up data from all students still enrolled in the state’s public schools. Other successful evaluations provided incentives to reduce comparison teacher attrition—monetary payments or promises that comparison teachers could receive professional development in the next program year.

Baseline Equivalence

Description. Experimental evaluations with high attrition, as well as all quasi-experimental studies, must establish baseline equivalence to demonstrate that no significant pre-intervention differences exist between treatment and comparison group participants on variables related to key outcomes. Establishing baseline equivalence ensures that groups have similar background characteristics.

Justification. Experimental evaluations with high attrition that demonstrate baseline equivalence of groups and quasi-experimental evaluations with demonstrated baseline equivalence of groups (or quasi-experimental studies where observed differences have been controlled for in analyses) are considered to be more rigorous. Baseline equivalence suggests that treatment and control groups were drawn from similar populations, thus making it less likely that differences between the groups

attributed to the interventions have alternative explanations or are due to confounding factors and biases.

Screening requirements. Experimental evaluations with high attrition and quasi-experimental evaluations meet the baseline equivalence criterion when their evaluation design meets at least one of the following two conditions:

1. There are no pre-intervention differences between groups on variables related to key outcomes that are greater than 5 percent of the pooled standard deviation.
2. Minimal pre-intervention differences exist between groups on variables related to key outcomes but those differences are controlled for in analyses.

Recommendations.

1. Report key baseline characteristics associated with outcomes for each group, such as pre-test scores and teaching experience. Always include sample sizes when reporting statistics.
2. Establish baseline equivalence using the exact sample included in the analyses of impacts. Thus, when reporting baseline equivalence, it would be helpful to only include those participants who are also included in the impact analyses in the tables and inference tests.
3. Conduct analyses on treatment and comparison groups that were comparable at baseline. Some successful evaluations began with data from a pool of potential comparison teachers who did not participate in MSP professional development. For their analysis, they then chose those comparison teachers who most closely matched treatment teachers on key characteristics. Successful evaluations matched treatment and comparison groups on such key characteristics as baseline test scores, school, district, grade level, teachers' years of experience and education, and ability level.

Assessing Validity and Reliability of Measurement Instruments

The third criterion requires that assessments and tests used to measure outcomes be valid and reliable. A measurement is considered reliable when it obtains similar results when retested or used by different raters. A measurement is considered valid when it has been shown to assess the outcome it was intended to assess. The same instrument should be used to measure outcomes in both the treatment and comparison groups

Quality of Measurement Instruments

Description. A crucial component of a rigorous evaluation design requires the use of high quality measures. This can be achieved through the use of existing data collection instruments deemed valid and reliable to measure key outcomes; instruments developed specifically for the study that are sufficiently pre-tested or have an internal consistency of 0.60 or higher; or data collection instruments composed of items from a validated and reliable instrument(s).

Justification. Evaluations must use instruments that accurately capture the intended outcomes and which have been tested on a group similar to the one being included in the study.

Screening requirements. All instruments used to measure outcomes must be deemed valid and reliable and have face validity (i.e., appear to measure what they purport to assess).

Recommendations.

1. Use instruments that have been shown to have accurate and consistent scores (i.e., have demonstrated reliability and validity). Where possible, use instruments that have demonstrated reliability and validity for a population similar to the population being studied. Successful evaluations used a variety of pre-existing assessments, including standardized state tests and test available on line in their subject areas.
2. Assessments created for the project must demonstrate validity and reliability using a population similar to respondents in the evaluation. For example, if the focus of the project is upper elementary school teachers, administer a pilot version of the assessment to 5th grade teachers in a school not participating in its program. The pilot results could then be used for assessing the reliability and validity of the instrument.
3. When selecting items from an existing measurement instrument:
 - a. Describe previous work that demonstrates that the scores are valid and reliable with a population similar to the current study;
 - b. Provide references to the manual or other studies discussing the validity and reliability of scores; and
 - c. Use full subscales rather than choosing items from across subscales where possible.

Assessing Whether Relevant Information is Reported

The fourth criterion requires projects to report relevant statistics so that readers can understand the effects of the intervention. The particular statistics that accomplish this varies depending on the design of the evaluation.

Relevant Statistics Reported

Description. The final component of our review required final reports to include treatment and comparison group post-test means and tests of statistical significance for key outcomes *or* sufficient information for calculation of statistical significance (e.g., mean, sample size, standard deviation/standard error).

Justification. Reporting relevant statistics provides critical context for interpreting the reported outcomes and indicates where an observed difference is larger than what would likely be created by chance.

Screening requirements. An evaluation meets the criterion if one of the following conditions is met:

1. Post-test means and test of significance for key outcomes are included in the evaluation.
2. The evaluation provides sufficient information to calculate statistical significance (e.g., reports of mean, sample size, standard deviations/standard error).
3. Other statistics are provided that indicates the significance and nature of the impact (e.g. effect sizes and impact estimates may substitute for post-test means and standard deviations/standard errors).

Recommendations.

1. For each evaluation, report means, standard deviations (or errors), and sample size. If reporting a regression model or ANOVA analysis, report the model statistics *and* means and standard deviations (or error).
2. Report the appropriate test for differences between groups. Successful evaluations reported both test statistics and significance values. For example, an evaluation with continuous gain scores on a standardized assessment reported t-tests and p-values for each of their findings. Another evaluation with a regression model of continuous outcome scores (controlling for baseline scores), reported coefficients and p-values. Those using ANOVA reported both the F-test statistic and the associated p-values.

Summary

As one of the goals of the MSP program is to assist projects in providing high-quality information on program outcomes, criteria were developed to guide projects in implementing and evaluating rigorous impact evaluations. These criteria are shared with all MSP projects and their evaluators and are described during annual regional meetings. Additionally, technical assistance to help projects meet the criteria is provided upon request.

While we recognize that not all projects are at the stage where rigorous designs are appropriate, particularly those that are still developing and testing hypotheses, the standards presented in the criteria are relevant to all evaluations, whether as guidance for future designs or for assessing current ones.

In the past six years, there have been substantial increases in the number of projects attempting to implement comparison group designs and in those implementing them successfully.

A summary of the criteria met in PP12 is helpful for understanding which elements of the criteria future projects may need additional guidance on when implementing their evaluations. Exhibit A.2 presents information on the number and percent of analytic contrasts that met each criterion, across all eligible evaluations. Since each criterion is not necessarily relevant for every analytic contrast, the total number of analytic contrast varies by criterion.²³

Exhibit A.2 indicates that evaluations were most likely to meet the criterion for *quality of measurement instruments* (87 percent of the analytic contrasts), followed by *relevant statistics reported* (65 percent of the analytic contrasts). Projects had more difficulty meeting the *attrition* criterion, with only one-third (33 percent) of relevant analytic contrasts meeting this criterion. This finding is consistent with anecdotal reports from projects that keeping teachers in the study and collecting post-test data from all participants is a challenge. Finally, only one-tenth of analytic contrasts (10 percent) met the criterion for *baseline equivalence*. This is not surprising given inherent differences between groups when randomization cannot be used. Further many projects, particularly those in rural locations, report having difficulty identifying appropriate comparison groups likely

²³ RCTs that meet the *attrition* criteria do not have to demonstrate *baseline equivalence*, and QEDs are only assessed on *attrition* if they only demonstrate baseline equivalence on the initial, and not on the analytic, sample.

leading to larger differences between treatment and control groups. Baseline equivalence is a difficult criterion to meet, and many well-designed evaluations do not meet this criterion.

Exhibit A.2: Number and Percent of Analytic Contrasts that Met Each Criterion for Rigorous Research Design, Performance Period 2012

Criterion (Number of Analytic Contrasts Relevant for Each Criterion)	Analytic Contrasts that Met Each Criterion	
	Number	Percent
Attrition (N=111)	37	33%
Baseline equivalence of groups (N=1397)	140	10%
Quality of the measurement instruments (N=1412)	123	87%
Relevant statistics reported (N=1412)	919	65%
Sources: Final evaluation reports, annual performance reports, and related documents submitted by MSP projects.		

Finally, as Exhibit A.3 indicates, the number of projects with at least one analytic contrast meeting all criteria increased four-fold from PP07 to PP09, and the proportion of projects with evaluations that meet all criteria has continued to increase

Exhibit A.3: Final Year Projects that Conducted Rigorous Evaluations and Met MSP Criteria for Rigor, Performance Periods 2007–2012

Projects	PP07	PP08	PP09	PP10	PP11	PP12
Implemented comparison group designs	37	49	65	59	59	71
Included at least one evaluation that met all criteria	4	3	16	15	17	21
Percent of projects with at least one evaluation that met all criteria	11%	6%	25%	25%	29%	30%

Local projects face many challenges in implementing rigorous designs, including such issues as limited resources, difficulties identifying reasonable comparison groups, and difficulties retaining all participants in the study and collecting their data. Additionally, local projects often lack evaluation expertise. Yet in an environment where there is greater attention being paid to the quality of research evidence, it has become increasingly important to support projects in implementing designs that are able to determine the effectiveness of their interventions. The MSP program has been educating its projects about rigorous evaluation designs by providing them with criteria for carrying out effective impact evaluations and, more recently, a user-friendly guide to the criteria (Bobronnikov, Sahni, Fernandes, & Bozzi, 2013).

Appendix B: Criteria for Classifying Designs of MSP Evaluations

This appendix includes the Criteria for Classifying Designs of MSP Evaluations used to determine the number of projects that successfully conducted rigorous evaluations. The criteria were developed as part of the Data Quality Initiative (DQI) through the Institute for Education Sciences (IES) at the U.S. Department of Education. The results of the review of final year MSP projects according to these criteria were presented in Appendix A.

Criteria for Classifying Designs of MSP Evaluations

- Experimental study**—the study measures the intervention’s effect by randomly assigning individuals (or other units, such as classrooms or schools) to a group that participated in the intervention, or to a control group that did not; and then compares post-intervention outcomes for the two groups
- Quasi-experimental study**—the study measures the intervention’s effect by comparing post-intervention outcomes for treatment participants with outcomes for a comparison group (that was not exposed to the intervention), chosen through methods other than random assignment. For example:
 - *Comparison-group study with equating*—a study in which statistical controls and/or matching techniques are used to make the treatment and comparison groups similar in their pre-intervention characteristics
 - *Regression-discontinuity study*—a study in which individuals (or other units, such as classrooms or schools) are assigned to treatment or comparison groups on the basis of a “cutoff” score on a pre-intervention non-dichotomous measure

Criteria for Assessing whether Experimental and Quasi-experimental Designs Were Conducted Successfully and Yielded Scientifically Valid Results

A. Attrition²⁴

- Met the criterion.** Key post-test outcomes were measured for at least 70 percent of the original sample (treatment and comparison groups combined) and differential attrition (i.e., difference between treatment group attrition and comparison group attrition) between groups was less than 15 percentage points.
- Did not meet the criterion.** Key post-test outcomes was measured for less than 70 percent of the original sample (treatment and comparison groups combined) and/or differential attrition (i.e., difference between treatment group attrition and comparison group attrition) between groups was 15 percentage points or higher.
- Not applicable.** This criterion was not applicable to quasi-experimental designs unless it was required for use in establishing baseline equivalence (see the *Baseline Equivalence of Groups* criterion below).

²⁴ The data reduction and baseline equivalent criteria were adapted from the What Works Clearinghouse standards (see http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf).

B. Baseline Equivalence of Groups

- Met the criterion (quasi-experimental studies).** There were no significant pre-intervention differences, as defined below, between treatment and comparison group participants in the analytic sample on the outcomes studied, or on variables related to the study's key outcomes. Two groups are considered to have baseline equivalence when:
 - the mean difference in the baseline measures was less than or equal to five percent of the pooled sample standard deviation; *or*
 - the mean difference in the baseline measures was more than five percent but less than or equal to twenty-five percent of the pooled sample standard deviation, and the differences were adjusted for in analyses (e.g., by controlling for the baseline measure); *or*
 - If the data required for establishing baseline equivalence in the analytic sample were missing (and there was evidence that equivalence was tested), then baseline equivalence could have been established in the baseline sample *providing the attrition criterion above was met.*
- Met the criterion (experimental evaluations that did not meet the attrition criterion above).** There were no significant pre-intervention differences, as defined above, between treatment and comparison group participants in the analytic sample on the outcomes studied, or on variables related to the study's key outcomes.
- Did not meet the criterion.** Baseline equivalence between groups in a quasi-experimental design was not established (i.e. one of the following conditions was met):
 - A. Baseline differences between groups exceeded the allowable limits; *or*
 - B. The statistical adjustments required to account for baseline differences were not conducted in analyses; *or*
 - C. Baseline equivalence was not examined or reported in a quasi-experimental evaluation (or an experimental evaluation that did not meet the attrition criterion above) and the necessary information was not provided such that reviewers could calculate it themselves.
- Not applicable.** This criterion was not applicable to experimental designs that met the attrition criterion above.

C. Quality of the Measurement Instruments

- Met the criterion**—the study used existing data collection instruments that had already been deemed valid and reliable to measure key outcomes; or data collection instruments developed specifically for the study were sufficiently pre-tested with subjects who were comparable to the study sample or the internal consistency of the instrument meets a minimum requirement of 0.60.
- Did not meet the criterion**—the key data collection instruments used in the evaluation lacked evidence of validity and reliability
- Did not address the criterion**

D. Relevant Statistics Reported

- Met the criterion**—the final report includes treatment and control group post-test means, and tests of statistical significance for key outcomes; or provides sufficient information for calculation of statistical significance (e.g., mean, sample size, standard deviation/standard error); or provides results from clearly specified statistical models.

- Did not meet the criterion**—the final report does not include treatment and control group post-test means, and/or tests of statistical significance for key outcomes; or provide sufficient information for calculation of statistical significance (e.g., mean, sample size, standard deviation/standard error); or provides results from clearly specified statistical models.

- Did not address the criterion**

Appendix C: 2012 State MSP Appropriations

MSP appropriations to states ranged from \$744,840 up to \$17,876,173, with an average of \$2,807,472 and a median of \$1,703,516.

Exhibit C.1: MSP Appropriations to the States

State	Total Funding Amount	State	Total Funding Amount
AK	\$744,840	MT	\$744,840
AL	\$2,685,939	NC	\$4,733,183
AR	\$1,611,191	ND	\$744,840
AZ	\$3,410,418	NE	\$744,840
CA	\$17,876,173	NH	\$744,840
CO	\$1,699,715	NJ	\$2,538,055
CT	\$887,579	NM	\$1,269,260
DC	\$744,840	NV	\$1,162,496
DE	\$744,840	NY	\$8,036,036
FL	\$8,120,268	OH	\$5,268,605
GA	\$5,260,953	OK	\$1,896,140
HI	\$744,840	OR	\$1,543,635
IA	\$940,605	PA	\$4,561,474
ID	\$744,840	PR	\$4,680,092
IL	\$5,282,424	RI	\$744,840
IN	\$2,899,118	SC	\$2,330,551
KS	\$1,048,080	SD	\$744,840
KY	\$2,210,718	TN	\$3,273,638
LA	\$2,550,206	TX	\$15,226,261
MA	\$1,707,316	UT	\$1,151,366
MD	\$1,470,908	VA	\$2,231,272
ME	\$744,840	VT	\$744,840
MI	\$4,663,571	WA	\$2,324,644
MN	\$1,603,345	WI	\$2,097,051
MO	\$2,412,599	WV	\$823,952
MS	\$2,071,960	WY	\$744,840