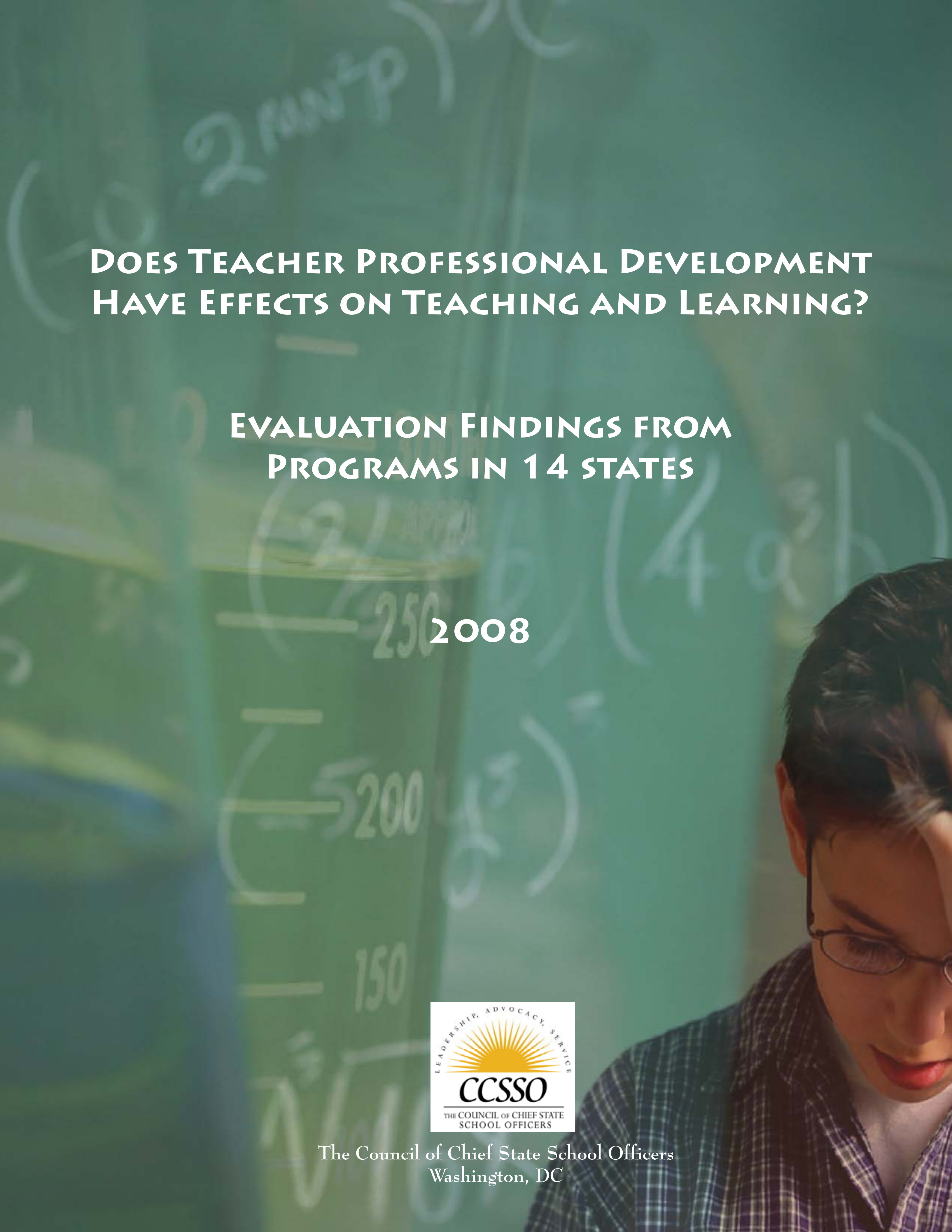


# Does Teacher Professional Development Have Effects on Teaching and Learning?

## Evaluation Findings from Programs in 14 states

2008

The Council of Chief State School Officers
Washington, DC

**Council of Chief State School Officers**
The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

**State Education Indicators**
The Council is a strong advocate for improving the quality and comparability of assessments and data systems to produce accurate indicators of the progress of our elementary and secondary schools. The CCSSO education indicators project is providing leadership in developing a system of state-by-state indicators of the condition of K–12 education. Indicators activities include collecting and reporting statistical indicators by state, tracking state policy changes, assisting with accountability systems, and conducting analyses of trends in education.

CCSSO also works with states on studies and reports to analyze effects of state education policies and programs.  This study is conducted under a grant to the Council of Chief State School Officers from National Science Foundation, Grant # REC 0438359.  This CCSSO study is also possible because of the excellent cooperation and coordination by staff in each participating state department of education.



Council of Chief State School Officers
Attn: Publications
One Massachusetts Ave., NW, Suite 700
Washington, DC 20001
202-336-7016
Fax: 202-408-8072
www.ccsso.org

ISBN: 1-884037-25-9

# Does Teacher Professional Development Have Effects on Teaching and Learning?

## Analysis of Evaluation Findings from Programs for Mathematics and Science Teachers in 14 states

Rolf K. Blank

Nina de las Alas

Carlise Smith

**February 2008**

Report prepared under a grant to the Council of Chief State School Officers from the National Science Foundation, Grant # REC 0438359:
**Project title: "Improving Evaluation of Professional Development in Mathematics and Science Education"**



**For further information, go to the Project Webpage**
http://www.ccsso.org/projects/improving_evaluation_of_professional_development

Council of Chief State School Officers, Washington, DC.

**Table of Contents**

**Does Teacher Professional Development Have Effects on Teaching and Learning? Analysis of Evaluation Findings from Programs for Mathematics and Science Teachers in 14 states**

In 2005 the Council of Chief State School Officers (CCSSO) began a study of teacher professional development programs in mathematics and science through a grant from the National Science Foundation. States nominated professional development programs for the study, and to conduct the study, the CCSSO team has worked with state coordinators and local program directors and evaluators. The study is designed to assist education leaders in all states by providing a cross-state analysis of the quality of professional development programs and evaluations using a common rubric developed from recent research on program effectiveness. For further information on the study design and results, see http://www.ccsso.org/projects/improving_evaluation_of_professional_development

## Summary of Findings

CCSSO reviewed evaluation studies from 25 professional development programs for teachers of mathematics and science from programs nominated by 14 states. The evaluation study reports and papers served as the data sources for the present analysis and paper. The reports primarily address evaluation findings from professional development activities conducted during the period 2004 through 2007. Following are several key findings from the study:

- **One-third of evaluation studies reported measurable effects of teacher professional development.** Seven of the evaluation studies of teacher professional development reviewed by CCSSO reported measurable effects of the teacher development activities on subsequent student outcomes. A total of 10 of the studies reported measurable effects on increasing teacher content knowledge, and four studies reported measurable effects on instructional practices of teachers.
- **Content focus plus sufficient time plus in-school component equals significant effects.** The cross-program review of studies showed significant effects of professional development programs for teachers of math and science when the programs include focus on content knowledge in the math and science subject areas plus training and follow-up pedagogical content knowledge. The total time in professional development for the studies with significant effects was 50 hours or more.
- **Purposeful evaluations yield measurable effects.** The evidence from the CCSSO review of evaluation studies shows that one-third of the programs reviewed had well-developed evaluations that produced findings with measurable effects on student achievement or change in instructional practices. Our analysis of evaluation findings emphasized scientific study design, and these kinds of designs could have been implemented across more programs.
- **Teacher vs. school-based professional development designs provide differing data on success.** Many designs for professional development are based on selection through teacher-based, voluntary models. The use of teacher-based professional development makes important follow-up activities harder to schedule and implement, and alignment to school curriculum more difficult to accomplish. For evaluation, the use of student assessment scores and tracking change over time appear to be facilitated with use of the school-based model for professional development.
- **Include outcome measures in allocation of evaluation resources.** Smaller professional development projects typically cannot afford ambitious, multi-stage evaluations or research. Allocation of more funds to evaluation would mean fewer participating teachers or fewer resources for the program implementation. In the cross-state program review we observe that smaller programs typically had to choose a few measures and methods of evaluation.

- **Plan for use of data systems and experimental designs.** Evaluations that will measure effects over time require access to data collection instruments or data systems, and advance planning with school officials. About one-fourth of program evaluations in the study did include comparison of a treatment group with a control group of teachers. State managers should consider evaluation designs that can be completed by linking data from state student assessments or local assessments with data on professional development for teachers.
- **Link teacher knowledge gains to change in classroom practices.** One type of evaluation finding identified in this review of studies showing promise for further use and expansion to other PD studies was measurement of change in teaching practices in the classroom. Four of the studies implemented well-tested instruments for comparing classroom practices across samples of teachers and classrooms. With advance planning, teachers and classrooms can be selected so that change in practices can be measured at the baseline point when teacher development begins and after implementation of activities and a period of implementation has been experienced.
- **Use findings in program decisions.** With the recent attention to scientific designs to provide measurement of impact of professional development on learning, we would like to see greater focus on how results from evaluations will be provided to decision-makers at specific points of time in the course of a project, and not long after the program activities have concluded. This is particularly important if a specific model is being considered for replication or expansion to other districts, schools, or additional teacher groups.
- **Value partnerships for evaluation.** Our analysis of evaluation findings across a number of programs and studies indicates that partnerships between higher education institutions and school districts have generally not added to the capacity for evaluation of professional development. For partnerships led by higher education institutions, the key partners with regard to data and measures for evaluation are local school district decision-makers and state education agency officials.

## Study Purpose: Analyze current teacher professional development in relation to recent research

In the present education policy environment, a high priority has been placed on improving teacher and teaching quality in U.S. schools. Standards-based educational improvement requires teachers to have deep knowledge of their subject and the pedagogy that is most effective for teaching the subject. States and school districts are charged with establishing and leading professional development programs, some with federal funding support, which will address major needs for improved preparation of teachers. Central to efforts to improve the quality of professional development is research-based evidence of effective programs and analysis of the characteristics of programs that make them effective.

Current policies at national and state levels are attempting to address the need for improving the preparation of teachers through professional development programs that have been demonstrated through research to be effective. Research over the past decade has provided a base of knowledge about the characteristics of effective programs of teacher professional development in mathematics and science. Recent federal funding directed toward teacher professional development, such as No Child Left Behind (NCLB) Title II and the National Science Foundation (NSF) teacher enhancement programs has reflected findings of research in the 1990s indicating the characteristics of programs that are effective in improving teaching and learning.

Effective math and science professional development has been found to have several common characteristics, including focus on content knowledge and skills of teachers, coherence with state and district standards for learning, sustained over time, and active methods of teacher learning of practices (see Banilower, Boyd, Pasley, & Weiss, 2006; Birman & Porter, 2002; Corcoran & Foley, 2003; Cohen & Ball, 1999; Cohen & Hill, 2001; Frechtling, 2001; Garet, Porter, Desimone, Birman, Herman & Yoon, 1999; Hiebert, 1999; Loucks-Horsley, Hewson, Love & Stiles, 1998; Kennedy, 1999; Weiss, Banilower, McMahon, & Smith, 2001). However, while a decade of research studies were finding strong evidence of what works, the data from large-scale national studies showed that most professional development provided to teachers did not meet these quality characteristics (Garet, Porter, Desimone, Birman & Yoon 2001; Desimone, Porter, Garet, Yoon & Birman, 2002; Corcoran & Foley, 2003). Thus in large part educators may know what kinds of programs should be developed and implemented based on research findings, but there are a variety of organizational, policy, and structural factors in education that have inhibited major change in practice.

Another emphasis of federal programs support since 2001 has been an emphasis on scientifically-based research and evaluation. State and local program grantees are asked to base their program designs for professional development on research evidence. Additionally, program designs are strongly encouraged to implement evaluation methods for professional development activities that use experimental designs, particularly randomized comparison trials that measure the outcomes of professional development using treatment and comparison groups. In many states and districts, program designers have worked to partner with researchers in universities, educational research organizations, or consultants to develop evaluations that are more robust, include validated measures and instruments, and track the effects of professional development efforts with teachers over time.

The CCSSO study on "**Improving Evaluation of Professional Development in Mathematics and Science Education**," supported by a grant from National Science Foundation, was designed to analyze the quality of a voluntary sample of professional development programs, as compared to the research, and to analyze the outcomes and evidence from the evaluation studies. Our overall goal is provide feedback to state education leaders about the degree to which this set of nominated professional development programs from a range of states meet criteria of high quality professional development established from research. We also wanted to identify common findings across programs on the effects of professional development on teachers and on their students.

**State Leader Needs.** State Education Agencies (SEAs) and state leaders have several roles in providing leadership with professional development for teachers. Each of these roles can benefit from research and evaluation findings concerning the program characteristics that have positive impact on improving teacher knowledge and skills and improving the quality of teaching and learning in classrooms:

- States write and set policies requiring teacher professional development, and in most states that have professional development requirements for re-certification the policies have a major effect on the types of professional development that are offered
- SEAs manage federal and state funds that are designated for teacher professional development and staff members often have responsibility for determining the size, scope, focus, and types of professional development teachers will receive, and many programs have specific evaluation requirements. Under the NCLB Title IIB funding for math and science states are responsible for awarding Math Science Partnership (MSP) local grants and each grantee must provide evaluations and reports on the effects of the MSP funding. As this program has grown since 2002, many states are taking a larger role in providing a statewide program evaluation design for grantees
- State education specialists who manage professional development or a curriculum subject area often provide recommendations to local districts and others concerning their professional development programs and strategies—including programs during scheduled in-service days, evening/weekend or summer programs, and teacher course credit options
- SEAs manage other areas that have significant and important implications for teacher professional development, including: development and implementation of state content standards for student learning, review and recommend curriculum materials and textbooks, and standards for teacher licensure and certification, and school improvement strategies and designs for low-performing schools and districts

**Use of the Study Results.** As a result of the important roles of state leaders in shaping and leading professional development in math and science education, CCSSO has undertaken the current cross-state study of the characteristics of high-quality, effective professional development. *Two key expected uses of the study results for state leaders are:*

- **Guidance on how to use findings from research on program quality—**in designing, selecting, or leading professional development programs supported by state agencies or in advising local agencies. State leaders will receive specific examples of the characteristics of professional development that will match what we know from research, and we can use evidence and examples from the current programs that are reviewed;
- **Assistance with evaluation designs, tools, and use of evaluations**—to improve the role of states in establishing designs and principles for evaluation and expectations for how the evaluation results will be used. All of the programs nominated by states for the CCSSO study have methods of evaluating effects of programs and activities, but the quality and appropriateness of methods used vary widely. The study provides a review of current methods and recommendations on how states can strengthen their evaluations and improve the usefulness of the evidence.

## Phase I: Findings on Quality of Professional Development

In the first phase of the study, CCSSO conducted a systematic review of 25 teacher professional development programs submitted by 14 participating states. The review was based on the program documents from proposals, designs, and initial reports. The review process was conducted through a PD Program Review Rubric and Guide, developed through the study, and teams of experts rated each of the 25 programs. The analysis addressed two questions:

- What is the quality of professional development across the nominated sample of programs, and what is the extent of variation in quality?
- What are the main program characteristics contributing to high ratings for quality that can be identified and replicated in future program design and development?

As a group, the programs could be considered representative of the current leading efforts to improve the teaching of math and science in public schools. The findings from the CCSSO study provide a way to analyze the status and prospects for math and science teacher professional development, and particularly initiatives supported through federal and state funding (see Blank, de las Alas, & Smith, 2007; http://www.ccsso.org/projects/improving_evaluation_of_professional_development).

The findings from first Phase program review can be summarized as follows:

**Content Focus:** The CCSSO analysis (completed at the end of 2006) found that current leading professional development is providing content knowledge development for teachers in math and science, especially for elementary and middle grades teachers. In 22 of 25 programs reviewed, the activities were rated as significantly focused on content knowledge in math or science. Additionally, a majority of programs reviewed were rated positively for providing important pedagogical content knowledge in math or science for teachers.
**Active Learning:** The professional development activities use active methods of learning for teachers in a large majority of programs. In comparison to the findings from research on professional development in the mid-1990s, the sample of programs in this study were surprising in the prevalence of active roles by teachers, including developing and presenting sample lessons, use of coaching and mentoring, developing new lessons or assessments, and interaction among teachers about ways to improve their practice.
**Collective Participation:** While most of the 25 reviewed programs did organize teachers and activities by common subject area and grade level, only a minority of the sample programs focused on delivery of professional development to teachers through a school-based strategy where teachers were learning with their school colleagues. The predominant organizing pattern was to plan teacher development for a treatment group drawn from a large number of schools and districts, with only a small number of teachers from each school.
**Coherence:** In almost all programs examined,, the reviewers found a description of how the program was designed to be aligned to state content standards. Additionally, a majority of program materials described how the development was consistent with local curriculum or with curriculum materials teachers were intended to use.
**Sufficient Time:** The average time for professional development activities including follow-up work in schools was found to be significantly greater than the typical math and science professional development documented and described in the mid-1990s.
**Evaluation:** The reviewed programs included evaluation designs with a number of evaluation objectives and tools. The CCSSO review covered four specific evaluation objectives, and a majority of programs included at least one measure in each objective: (a) quality of implementation of development activities; (b) gains in teacher knowledge; (c) change in classroom practices, and (d) increase in student achievement. Thus, in general, the programs were very ambitious in emphasizing methods of evaluation.

## Phase II: Analysis of Evaluation Results across States

The second phase of the study, the subject of the present report, is an analysis of outcomes and findings from the program evaluation reports. CCSSO has analyzed data and findings from evaluations conducted for the 25 programs in our study. This second study report addresses the following four questions about the results of our analysis of evaluations of professional development:

- What evaluations were completed and what findings were reported?
- What were the types of major findings from the evaluations? How were they measured? What measures of outcomes were used in the PD evaluations?
- What conclusions can be drawn about the adequacy and usefulness of the evaluations and reports? What are the cross-report recommendations that are useful to state leaders and evaluators?

CCSSO has compiled and analyzed the evaluation studies and reports completed and provided us as of September 2007. Thus the period of evaluation findings that were included in our cross-state analysis were prepared and reported from Spring 2005 through Spring 2007.

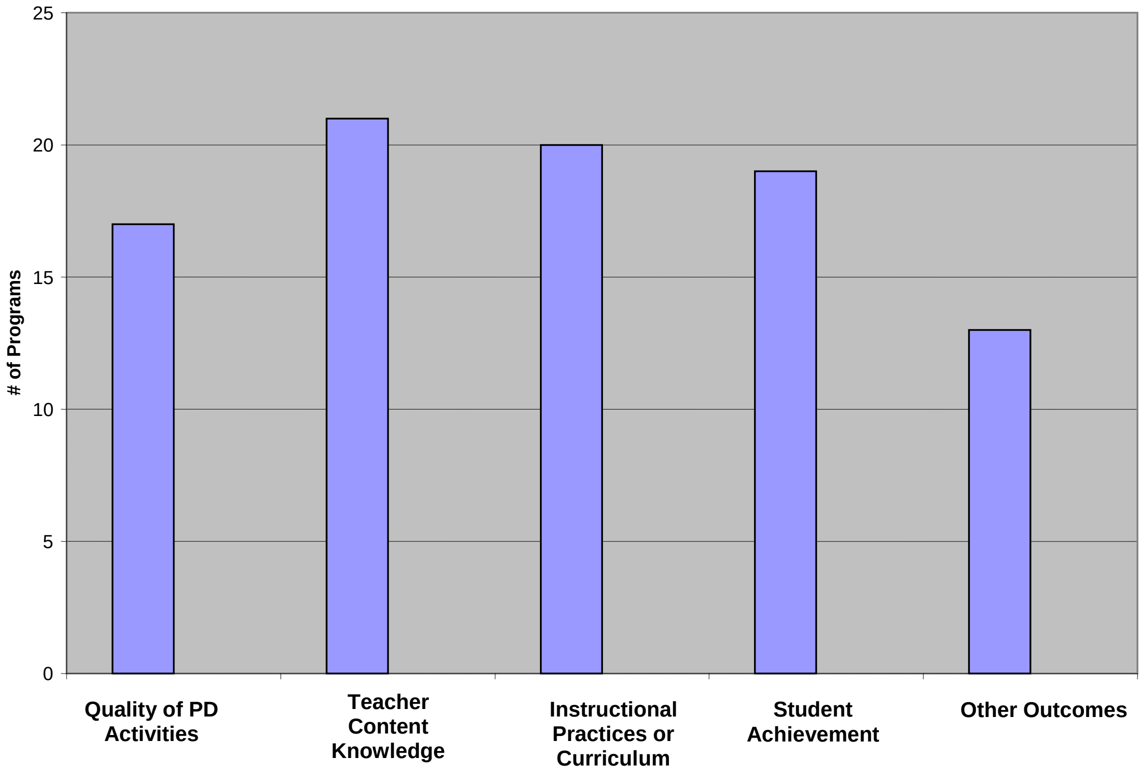
Study Evaluations Completed and Reported

**Initial Evaluation Designs.** In the first phase of the CCSSO study, we analyzed the designs for evaluation of professional development and we categorized the types of evaluations intended to be used. The 25 programs in our study were identified in Spring 2005, and the descriptions of intended professional development activities and evaluation designs were reviewed by CCSSO in 2005 and 2006. The analysis of evaluation designs of the teacher professional development programs was conducted using four categories of outcomes. These categories are consistent with the logic model for the CCSSO study design (CCSSO, 2005) based on our review of the research:

(a) quality of implementation of professional development activities
(b) gain in content knowledge and pedagogy skills of teachers
(c) change in instructional practices or curriculum that is taught
(d) improvement in student achievement

As shown in Figure 1, almost all of the programs had planned to carry out evaluations intended to assess professional development across at least three of the four categories of outcomes. Seven had evaluation designs that cover all four categories and other outcomes. These programs utilize multiple measures to report to their stakeholders on how the program is being implemented and the degree to which the program activities are having effects on teaching and learning. In several cases, the programs also track other outcomes specific to the program.

**Figure 1 - Summary of Intended Program Evaluation Design**




**Evaluation Reports and Studies.** In Fall 2006, CCSSO requested evaluation reports from all of the sample programs that included specific evaluation designs for professional development, and the information was summarized above. In Spring 2007, CCSSO communicated a second request for evaluation reports and studies to all of the participating program directors and evaluators.

As of September 2007, CCSSO received a total 41 evaluation reports or studies, and thus a majority of the programs provided multiple reports.  These documents served as the data sources for the present analysis and paper. The reports address evaluation findings from professional development activities conducted during the period 2004 through 2007. Reports vary in analyzing data for one, two, or three years of activities. Table 1 lists the evaluation reports received across the programs in our study. The columns of the table indicate the type of data reported by evaluation category (e.g., PD activities—measured by teacher survey). The column on the far right indicates the number of teachers that involved in the professional development and their subject and grade level.

The 41 reports include a range of outcomes data and descriptive analyses, and some of the evaluation evidence pertains to program outcomes other than for professional development, including pre-service, curriculum materials, and partnership development. Many of the evaluation reports included formative, descriptive data which were intended as internal evaluation information primarily used by the study team and institutional administrators. In the table, the CCSSO study team identified the data and findings from four types of evaluation outcomes from professional development. These data and findings were the focus of our cross-state and cross-study analysis.

**Table 1: Program Evaluation Reports by Objective**

| Program | Report Date | Student Achievement | PD Activities | Teacher Knowledge | Instructional Practices | # of Teachers |
|---|---|---|---|---|---|---|
| Northeast Front Range (CO) | • November 2004 - Yr 1 Report<br>• March 2006 - Yr 2 Report (Jan.–Dec. 2005)<br>• March 2007 - Yr 3 Report (Jan.–Dec. 2006) | • Student achievement linked to treatment teachers<br>• Change from 2004 to 2005, 2005 to 2006<br>• Treatment vs. Control Groups | • Measured through teacher survey by course | • Measured through pre-post test by course | • Measured through teach survey on student<br>• Examined impact through focus group | • 136 treatment teachers/yr. |
| Rocky Mountain Middle School (CO) | • May 2006 - Yr 2 Report<br>• May 2007 - Yr 3 Report | • No student achievement scores reported<br>• Students with advanced courses by district | • Measured through teacher surveys<br>• Measured through survey of IHE staff | • Measured through teacher knowledge test<br>• Analyzed pre-post gains per course reported | • Measured through teacher survey on instructional practices | • 175 math & science teachers |
| Teacher Quality Enhancement/Secondary Teacher Enhancement Project - TQE/STEP CO) | • December 2005 - Year 1 Report<br>• October 2006 - Year 2 Report | • Student achievement by year for student of new Gr. 6-10 teachers<br>• Reported on Baseline Year - 2005 | • Measured through survey of new teachers<br>• Retention rate of new teachers/yr. | • Measured through PRAXIS, PLACE<br>• Measured through exams of new teachers | - None - | • 364 teacher leaders |
| Science Lead Teacher (DE) | • Report, no date | - None - | - None - | - None - | - None - | - None - |
| Science Coalition Courses (DE) | • October 2000 | - None - | - None - | - None - | - None - | - None - |
| Launch II (FL) | • 2006<br>• September 2006 (External Evaluator) | • Reported on 2005 to 2006 school mean in math and in science vs. district mean | • Measured through survey of teachers on PD activities | • Measured through pre-post teacher knowledge test | • Measured through survey of instructional practices | • 161 elementary & middle teachers |
| Developing Mathematical Thinking - DMT (ID) | • 2005<br>• 2006 | • ISAT planned<br>• Direct Math Assessment (K-2) | - None - | • Measured through teacher math knowledge inventory<br>• Analyzed pre-post gains | • Measured through SEC surveys on instructional practices | • 60 elementary teachers |
| iCATS: Beyond the Textbook (IN) | • May 2004 - Year 1 Report<br>• 2005 - Year 2 Report<br>• July 2006 - Year 3 Report | • Achievement gains of students of treatment teachers vs. gains of students of control teachers<br>• Reported pre-post tests for Gr. 4 & 5, Year 1 & Year 2, Gr. 3, 4, & 5 in Year 3 | - None - | - None - | - None - | • Approx. 57 teachers/yr. |
| Making Math More Meaningful – M4 (IN) | • 2005-06 report | • Reported change in ISTEP student achievement (Grades 3-8) from 2003-04 to 2004-05 of students from treatment math teachers, no comparison group | • Based on interviews w/ teachers<br>• PD is associated with the implementation of *Everyday Math* curriculum | - None - | • Based on observations of 26 teachers | • Elementary teachers: approx. 36 per summer institute, 57 per in-school sessions |

**Table 1 – continued**

| Program | Report date | Student Achievement | PD Activities | Teacher Knowledge | Instructional Practices | # of Teachers |
|---|---|---|---|---|---|---|
| High School Mathematics Alliance Grant (KY) | • No report | - None - | - None - | - None - | - None - | - None - |
| Math/Science Teachers Engaged as Mentors - MS*TEAMS (KY) | • 2006-Year 2 Report | - None - | • Descriptive report of evaluator<br>• Interviews with teachers | - None - | - None - | • 21 teachers<br>• Learning teams |
| Mathematics Access and Teaching in High Schools - MATHS (ME) | • September 2005 - Year 1 Report<br>• August 2006 (2004-06 implementation period) | • Measured through Gr. 9 Algebra Concepts Test<br>• Student interviews & focus groups, classroom observation<br>• Linked to treatment teachers<br>• Analyzing student achievement change from 2005 to 2007 | - None - | - None - | • Based on classroom observation, analyzing types of instruction<br>• Level of cognitive activity<br>• Reported through student focus groups | • 100 middle school & high school math teachers |
| EduTron (MA) | • 2006<br>• Fall 2006 Addendum | • Student achievement district level change from 2003 to 2006 | • Pre-treatment assessments of participant teachers determined type of course offered | • Measured through MA teacher licensing test (MTEL)<br>• Pre-post teacher knowledge test | - None - | • 127 elem/mid school math teachers in 3 districts<br>• 40 elementary & middle school math teachers in last course cycle (Dec.-Apr.) |
| Coalition for Higher Standards Math Partnership Program - Lesley/MassInsight (MA) | • 2006<br>• January 2006 (External Evaluator) | • Reported on student achievement for 2004, baseline year<br>• Math teachers vs. control teachers, by school | - None - | • Measured through teacher knowledge test<br>• Treatment teachers vs. control teachers | • Based on interviews of coaches, teachers<br>• Based on classroom observations | • 100 elementary & middle school math teachers as coaches in 4 school districts |
| Teachers as Leaders and Learners - TaLL (NJ) | • October 2005<br>• October 2006 | • Student achievement linked to teachers<br>• Analyzing change from 2004 to 2005, and 2005 to 2006 in Terra Nova (dist test)<br>• Treatment vs. control groups | • Measured through teacher survey on practices and attitudes | • Based on grades in math courses | • Measures on teacher survey<br>• Grades | • 20 middle school math teachers treatment group |
| Consortium for New Explorations in Coherent Teacher Education - CONNECT-ED (NJ) | • October 2004, pilot project | - None - | • Based on interviews, observations | - None - | - None - | • 40 treatment teachers in 10 design teams |
| Pursuing Excellence in Middle School Math & Science – Nash-Rocky Mount (NC) | • July 2005 Report<br>• October 2006 [Data table]<br>• October 2006 (2004-05) (External Evaluator)<br>• October 2006 final report (2004-05)<br>• January 2007 report | • Change in % of students at proficiency or beyond by grade level (grades 6-8) from 2003-04 to 2004-05 in end-of-grade math assessment<br>• % of students at proficiency in 2005-06 (preliminary data) for Gr. 6 in end-of-grade and alternative math assessments | • Based on teacher focus group interviews | • Analysis of teacher assignments for increased math understanding using 5 levels<br>• PRAXIS | - None - | • 35 teachers and coaches all levels |

**Table 1 – continued**

| Program | Report date | Student Achievement | PD Activities | Teacher Knowledge | Instructional Practices | # of Teachers |
|---|---|---|---|---|---|---|
| BreakThrough Mathematics (OH) | • June 2006 | - None - | • Measured through teacher follow-up survey on perceptions, satisfaction on program | - None - | • Measured through survey on applying newly acquired knowledge/skills to practices | • 194 teachers taking online math course |
| Comprehensive, Data-Based Professional Development for Middle School Mathematics Teachers (OH) | • May 2006 | • Percent proficient by school in math and/or science, baseline from 2004-05 results | • Based on evaluators/program implementers' observations (no protocol) | - None - | • Classroom observations and interviews by coaches | - None - |
| High Desert (OR) | • 2005 | • Reported on student achievement change and comparing treatment vs. control groups, based on results in 2004 to 2005 | • Teacher participation rate | • Measured treatment teacher knowledge through DTAMS<br>• Compared PRAXIS results between treatment vs. control teachers | • Measured through SEC, RTOP<br><br>• Needs assessment | • 24 elementary math teachers |
| Willamette Valley Watershed Project (OR) | • September 2005<br><br>• October 2006<br><br>• October 2006 [Part 2: Narrative Responses] | • Reported on student achievement in science based on test specific to the project.<br>• Compared treatment & control group, using pre-post testing | - None - | • Measured through teacher knowledge test linked to student achievement test | • Analyzed lesson plans | • 15 treatment teachers, 15 control teachers |
| Oregon Mathematics Leadership Institute - OMLI (OR) | • May 2006<br><br>• April 2007 | • Reported on % of students who met or exceeded the math standards on state assessments between 2004-05 & 2005-06<br>• Reported on results from student discourse observation | - None - | • Measured through teacher survey<br><br>• Measured through teacher content knowledge test, reported pre-post results | • Measured through instructional practice survey on sample of teachers<br><br>• Based on classroom observation | • 280 School Leadership Team members (2 teachers & 1 administrator per team) |
| Mathematics and Science Coaching Initiative – South Carolina Coaching Initiative (SC) | • 2007 report on Fall 2003-Spring 2006 | • No student achievement data reported | - None - | - None - | • Measured through teacher survey to track instructional change | • 33 science and math teacher coaches<br><br>• 600+ teachers |
| Northwest Wisconsin Partnership for Mathematics and Science Education – NW Wisconsin (WI) | • December 2005<br><br>• n.d. | • No student achievement data reported | • Measured through teacher survey<br><br>• Measured through interviews | • Measured through teacher self-rating of knowledge gains | - None - | • 30 middle school math and science teachers |
| Wisconsin Academy Staff Development Initiative Retention and Renewal – R & R (WI) | • June 2005, 2004-05-Year 2 Exec. Summary | • No student achievement data reported | - None - | - None - | - None - | • 40 teachers/year<br><br>• Teacher mentors |
| **Number of Evaluation Reports** | **41 total** | **20 reports with student achievement data**<br>**11 reports with treatment vs. control comparison** | **14 on PD activities** | **24 on teacher knowledge** | **13 on instructional practices** | |

**Implementation of PD. A total of** 14 reports included evaluation data on the quality of how professional development activities were implemented (i.e., extent of teacher participation, fidelity of implementation of PD design, and response of teachers to the professional development). All of the evaluation reports include reports on how the activities were implemented and descriptions of the activities. Some of the evaluation studies heavily focused on the use of information as formative evaluations to assist decision-makers and leaders to assess how the program is operating and how improvements can be made. However, some of the data provide outcome findings on the quality of implementation of the professional development.

**Gains in teacher knowledge.** Twenty-four of the 41 evaluation reports included data and findings regarding gains in teacher knowledge related to the professional development activities being evaluated. Further analysis of these data across sites will allow us to determine whether there are common factors in producing gains in knowledge and how the several knowledge tests worked to measure gains.

**Change in classroom instructional practices related to the professional development.** Thirteen reports included data on instructional practices of the teachers participating in professional development, allowing program evaluators to determine if there was change in instructional practices of participating teachers. A variety of measures and tools were found to be in use.

**Improvement in student achievement related to teacher professional development.** Twenty of the evaluation reports included data on student achievement for teachers, schools, and districts in which teachers were involved with the professional development activities. A major question for our CCSSO study was the degree to which student achievement gains could be attributed to the treatment of professional development. Sixteen evaluation reports included student achievement trends for at least two years. A small number tracked achievement gains for more than two years.

In sum, the evaluation reports do not cover as many of the design objectives as were outlined and intended in the program proposals and initial descriptions. The reports received include findings from slightly more than half of the programs in any one category, or significantly fewer results than originally planned. It is possible that additional evaluation reports and studies will be released by programs.

## Analysis of Major Findings Identified in Evaluation Reports

CCSSO staff reviewed the findings from the evaluation reports submitted from the programs in the study. The goal of CCSSO's analysis was to identify findings from the reports that are based on measurable effects of teacher professional development. A set of criteria was established for determining measurable effects under each of the evaluation categories. This approach to identifying cross-state findings allows CCSSO to highlight evaluation studies that can be held as examples of professional development that provide clear, scientific evidence of impact on teaching and learning.

The studies for which we found measurable effects of professional development are grouped by type of outcome. We provide the core evidence in a table that allows the reader to review findings across studies and to understand the basis for selection of measurable effects.

## Student Outcomes

We display in Table 2 the findings of effects of teacher professional development on student outcomes. Three primary criteria were used to determine findings that can be classified as demonstrating effects:

a) Finding of effect on student outcome is supported by statistical significance of change linked to the treatment teachers
b) Finding is substantively important, i.e., an educationally significant change
c) Measure of student outcomes is reliable and valid for the evaluation purpose

**Table 2: Evaluations Reporting Measurable Effects of Teacher Professional Development on Student Outcomes**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| META Associates. (2007), **NE Front Range**-CO | Student gains on CSAP Math for two years | Quasi-experimental Pre-/posttests | Matched comparison group of teachers | CSAP state assessment results show statistical significant mean gain of 13.5 points for students of participant teachers, with 70% improved in scale score, and higher gains than students of comparison teachers (higher initial scores). |
| Schmidt, D. L. (2006), **Launch II**-FL | Student gains on FCAT elementary math, science assessment | Quasi-Experimental Pre-/posttests | Teacher PD whole school vs. weighted district level performance | FCAT results show statistically significant gains on FCAT math in Grades 4-5 for treatment schools (+2.1%) vs. comparison (+.8%). |
| Evansville-Vanderburgh Schools. (2006), **iCATS**-IN | ICAT math gains for Grades 3, 4, and 5 for 2 years | Quasi-experimental Pre-/posttests | Treatment students vs. students participated in direct teaching environment | Statistically significant gains on ICAT Math for Grades 3 – 5 students of teachers receiving PD with Educational Technology (gain scores 4 points higher than comparison students). |
| Grip, R. S. (2006), **TaLL**-NJ | Gains on district-wide Terra Nova test in math & science | Pre-/posttests School-by-school analysis | Treatment vs. non-treatment teachers in same school | Over two years, significant gains on Terra Nova test for students of teachers in PD (in math 7-pt. higher gain; in science 2-pt. higher gain). |
| Hansen, J. B. (2006), **Willamette Valley**-OR | Gains on Student Science Achievement Level Test (developed with Northwest Evaluation Assoc.) | Quasi-experimental Pre-/posttests | Matched comparison group of teachers | Treatment group outperformed the comparison group at grades 8, 9, but not grades 6, 7. Effect size moderate (.106) across grades. |
| Weaver, D. (2007), **OMLI**-OR | Change in student discourse as observed thru teacher surveys, Student Discourse Protocol | Quasi-experimental Pre-/posttests | Teachers in treatment matched to non-participant comparisons | Student discourse in classes of teachers in treatment group significantly greater than comparison teacher classes. |
| Niess, M. L. (2005), **High Desert**-OR | Student gains on state math assessment for elementary & middle grades | Quasi-experimental Pre-/posttests | Matched comparison group of teachers | From 2004 to 2005 elementary treatment group showed statistical gains in math achievement (214 to 220, p<.05) vs. comparison group (210 to 217 p<.05). |

**Measurable effects on student outcomes (7 studies).** The seven studies in Table 2 were identified in the CCSSO review as reporting at least one measurable effect of the professional development initiative on improving student outcomes. Although the size of the effects varies, there are common characteristics. First, the evaluation design allowed the evaluator to have comparison of data so that differences could be statistically tested. In five studies, the student outcomes data were tracked over time using student achievement test data, and then differences in achievement scores could be statistically tested for change at the second point. The measurement of change over time was not based on individual student records and test scores, but the design did involve matching teachers in the treatment group with the student achievement of their students and then tracking the student achievement for students in the same grade during the next year or succeeding years.

A second characteristic of these studies showing effects is a quasi-experimental design. Random assignment of teachers into groups was not used (to produce true treatment vs. comparison groups), but each study was able to identify a comparison group of teachers that matched the characteristics of the treatment group (such as same grade level, subject, district, student population). The comparison groups in the six separate studies were obtained in different ways. One study matched treatment schools with comparison schools that did not have the program, another study used a weighted sample of teachers drawn from the same district as the treatment group teachers, and a third study matched individual treatment teachers with comparison teachers. By analyzing student outcomes for teachers in the treatment group against a comparison group of teachers, study readers and users have greater confidence in the claim of improved learning due to the professional development. In addition, statistical methods can be applied to scientifically determine if the differences claimed are not merely by chance.

The specific measure that is used to determine student learning also is a key characteristic that can increase the confidence and validity of the finding. Five of the studies with measurable effects used data from statewide student assessment programs to track change by year. One study used a teacher survey and class observation protocol to analyze student discourse in class. The statewide assessments have a major advantage as an outcome measure by providing an opportunity for longitudinal analysis by teacher that does not require additional data collection or test design. Also, student performance can be tracked from year to year, or cohort grade level data can be compared from year to year. A possible disadvantage is the comprehensive nature of the assessments—with the primary purpose of accountability at the school level. The assessment may not be sensitive to the content area(s) or pedagogical knowledge emphasized in the teacher professional development. The Weaver study employed an outcome measure of student discourse that was designed to be a direct measure of an intended outcome of the professional development curriculum.

**Table 3: Evaluations with Findings on Student Outcomes but not Measurable Effects**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| Heath, B. P, et al. (2007), **Rocky Mountain MS**-CO | Change in student attitudes about math and/or science after teacher PD | Case study with pre-/post-treatment attitude surveys of student camp participants | No comparison group | Student attitudes about math and science courses more positive after teacher professional development, but small N-sizes. Attitudes are indirect measure of learning gains. |
| JVA Consulting, LLC. (2006), **TQE-STEP**-CO | Change in student achievement in grades 6-10 on state assessment scores across 4 subjects areas (math, science, reading, writing) | Analysis of scores of students of new teachers in SY2005-2006 compared to scores of students of teachers hired before SY2005-06 | No comparison group – sets baseline prior to treatment teachers entering classrooms | Shows students of teachers hired before 2005-06 to have nearly the same scores compared to students of new teachers. |
| Howard, M. N. (2006), **Launch II**-FL | Changes in percent of students proficient in state math assessment; changes in mean scale scores in grade 5 and 8 state science assessment | Analysis of student achievement results in state math and science assessments aggregated to participating schools compared to districts | No comparison group | Students of teacher participants have the same level of improvement as district overall average. Limitations of results include: masking actual participant teachers' impact on student achievement due to PD, math and science results are not comparable, and science effects of participant teachers not in grade 5 or 8 not measured. |
| Rosenblum, J. (2006), **MATHS**-ME | Change in student engagement and cognitive activity, change in student attitude; change in a Gr. 9 Algebra Concepts Assessment | Mixed-method with class observations, student focus group interviews, and analysis of results from Gr. 9 Algebra Concepts Assessment | No comparison group | Increased range in student engagement and cognitive activity after teacher PD and improved attitudes toward math course; Slight improvement in mean scores on one strand of the assessment. |
| Chen, A. (2006), **Edutron**-MA | Changes in percent of students proficient in state math assessment at each school levels | Analysis of change by year in participating districts' student proficiency percentages in math | No tracking of treatment group | Student proficiency percentages increased at all 3 school levels but cannot be attributed to Math PD for elementary teachers. |
| Perry, M. (2005), **Nash-Rocky Mount**-NC | Change in percentage of students proficient in end-of-grade state tests for the district | Analysis of change by year and grade (6, 7, 8) in participant districts' student proficiency percentages in math | No tracking of treatment group | Aggregated at the district level, percentage of students deemed proficient increased on average by about two percentage points for middle grades. |

**Student outcomes with no measurable effects**. From the CCSSO review of the evaluation reports, we found that six additional evaluation studies reporting findings on student achievement related to professional development. However, the findings in these studies did not meet our study criteria of demonstrating measurable effects, as outlined above.

In three of the studies submitted by JVA Consulting, Chen and Perry, only district level achievement scores were reported for the districts in which teacher professional development was provided. Even though the data allow pre-post analysis, these results do not allow the evaluator to determine the effects of professional development treatment from many other variables that could produce change in student achievement. The students that might be affected by teachers experiencing the program were not distinguished from students whose teachers were not receiving teacher PD. In these studies, the evaluator did not have sufficient access or comparison of the data for teachers and students to be able to test the question of treatment for teachers and effects on outcomes fro their students.

Two studies reported student attitude change as a result of professional development. Although attitudes can be measured as an outcome, and may be correlated with improved learning over time, attitude data provide only an indirect measure of student learning gain. Also neither of the two studies had a comparison group and the number of teachers and students studied were small. Another study reported survey responses from teachers regarding student achievement, and while the responses indicated gains in achievement this measure provides only indirect evidence of an outcome linked to teacher professional development.

## Teacher Content Knowledge

**Effects of professional development on teacher knowledge outcomes (10 studies).** A total of ten of the evaluation reports included at least one finding that demonstrated a measurable effect of professional development on increased teacher knowledge in math or science. The findings reported in Table 4 were selected using two criteria for determining measurable effects of teacher knowledge were used across these studies:

a) Assessment of knowledge gain based on an instrument that was previously validated, or
b) Assessment of knowledge that was mandated by the state for teacher licensure/certification.

All the studies identified as having measurable effects were able to determine gains through pre- and post-testing of teachers in the treatment group. Only one study employed a quasi-experimental design. Apparently, most of the program evaluators assumed that it was sufficient to measure only gains in knowledge for the treatment group, and that documenting the gains could be directly attributable to the program using the pre-post model (without testing a comparison group of similar teachers). Testing of teachers that do not have a commitment to the activity (e.g., in a comparison group situation) can be difficult, however, a comparison group of teachers not in the PD or possibly in a somewhat different type of professional development would provide useful comparative data.

**Table 4: Teacher Knowledge Effects: Evaluations Reporting Measurable Effects of Teacher Professional Development on Teacher Content Knowledge**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| META Associates. (2007), **NE Front Range**-CO | Gains in teacher knowledge as measured by Content-Based Teacher Assessment | Analysis of pre-/posttests given before/after summer institutes | Treatment teachers only | Statistically significant gains by teachers on biology assessment (N=21, 25% gain <.001), earth/space (N=17, 26% gain, <.001) force & motion (N=25, 22.3% gain, <.001. |
| Heath, et al. (2007), **Rocky Mountain MS**-CO | Impact of courses as measured by teacher gains on Teacher Content Inventory | Analysis of pre-/posttests given before/after summer institutes | Treatment teachers only | 12 of 15 academy courses with medium to large effect sizes on teacher knowledge: Earth Science II (3.8), History of Math (3.35), Physics II (2.23), Earth Science I (1.89), Discrete Math (1.69), Math Modeling (1.68), Chemistry I (1.63), Physics I (1.5), Biology I (1.41), Biology II (1.25), Geometry (1.18), and Statistics (.7). |
| Schmidt, D. L. (2006), **Launch II**-FL | Gains in teacher knowledge as measured by Teacher Content Knowledge Assessment | Analysis of pre-/posttests given before/after summer institutes | Treatment teachers only | Substantive increase in grades K-5 teacher scores in PD program for three cohorts: Cohort I (2004-06, N=10, 16% pt. mean gain), Cohort II (2005-06, N=36, 27% pt. mean gain, and Cohort III (N=50, 18% pt. mean gain. pre/post 2006 summer institute). |
| Brendefur, J., et al. (2005), **Developing Math Thinking**-ID | Gains in teacher knowledge as measured by Number Knowledge Inventory | Analysis of pre-/posttests given before/after treatment | Treatment teachers only | 20% gain by teachers in number knowledge was statistically significant (Pretest mean score=31.35; Posttest mean score=40.49, $p < .01$). |
| Chen, A. (2006), **Edutron**-MA | Number of teachers passing the Massachusetts Tests for Educator Licensure (MTEL); Gains in teacher knowledge as measured by course assessments | Analysis of posttest only treatment (MTEL); Analysis of pre-/posttests given before/after courses | Treatment teachers only | 14 of participant teachers passing MTEL; teachers showed knowledge gains in geometry (N=19, 22.6 pt. gain from 61.3; and probability (N=19, 34.1 pt. gain from 48.1). |
| Lesley University. (2006), **Coalition for higher math standards**-MA | Gains in course assessments | Analysis of pre-/posttests of courses | Treatment teachers only | 19 of 30 K-5 teachers, 24 of 40 middle school teachers and 5 of 7 high school teachers showed statistically significant ($p<.05$ level) increase in knowledge of mathematics. |
| Hankerson, L. (2006), **Nash Rocky Mount**-NC | # of teachers passing PRAXIS, thus earning HQ status | Analysis of pre-/posttests | Treatment teachers only | Number of non-HQ math and science teachers who reached HQ status increased from 80 to 101 and from 101 to 105, respectively. |

**Table 4 – continued**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| Niess, M. L. (2005), **High Desert**-OR | Gains in DTAMS, PRAXIS | Quasi-experimental Analysis of pre-/posttests of DTAMS, Analysis across 2 teacher groups | Treatment and comparison groups (PRAXIS) | Treatment group increased math knowledge on DTAMS (N=24, pretest mean=76.9167, posttest mean=97.7500, SD=19.49861 significant at p<.05 level). In PRAXIS, there is a statistical difference between the two groups on the problem solving domain, with treatment (N=14) group teachers scoring significantly better than their comparison (N=13) counterparts. (T=2.848, significant at p<.05). |
| Hansen, J. B. (2006), **Willamette Valley**-OR | Gains in knowledge assessment specific to PD; Changes in lesson plans content as seen thru Chief State Science Supervisors rubric | Quasi-experimental Pre-/posttests | Treatment and Non-equivalent comparison group (self-selected) | Statistically significant gains for treatment teacher group (pretest mean=73.0, gain=6% pts.) vs. comparison group; lesson plan analysis showed statistical difference (p<.05 level) in treatment & comparison groups: *proposing explanations* (.p=.006), *alternative explanations* (p=.002), *linking explanations (*p=.016), *communication* (p=.024). |
| Weaver, D. (2007), **Math Leader Institute**-OR | Gains in U. of Michigan Content Knowledge Assessment (CKT-M) | Analysis of CKT-M results | Treatment teachers only | Secondary school treatment teachers showed statistically significant math gains (N=81, mean difference=.025), and elementary teachers showed significant math gains (N=93, mean difference=.037). |

A recent emphasis of program funding agencies for math and science professional development is assessment of teacher knowledge. The National Science Foundation has strongly recommended that MSP grantees employ assessments of teacher knowledge, and the guidelines of U.S. Department of Education for Title IIB MSP grants also emphasize use of teacher assessments. These program guidelines follow from the research consensus findings at the end of the 1990s that content knowledge focus is a critical element to effective professional development.

Two studies (Chen, 2006; Hankerson, 2006) reported on the number of participant teachers achieving certification and highly qualified status as a result of passing state teacher assessments. The assessment was a part of the state licensure process, and the programs were measuring gains in teacher certification as an outcome.

The programs in this study analyzed several teacher assessments that are commonly employed across professional development in other states. The evaluations of META Associates (Front Range) and Heath, et al. (Rocky Mountain) had knowledge assessments specific to each subject area course—geometry, chemistry, algebra, earth science, etc.—that were offered to teachers. The assessments were given immediately prior to the course and at the end of the course, typically taught in summer institutes, and almost all the courses had high rates of increase in teacher knowledge. The study by Niess (High Desert) used the DTAMS assessment and results showed that teachers gained significantly in four math content areas assessed. Weaver (OMLI) used the University of Michigan content assessment in mathematics and documented gains for elementary and secondary teachers with the gains differing by content scale. The design for the Weaver study also tested knowledge gains for a comparison group of teachers matched to the treatment group, and found significant differences in gains for the treatment group. The Willamette Valley project developed a science teacher assessment that was specific to the goals of the professional development and found significant gains. The study also evaluated change in lesson plans as a second measure of teacher knowledge, and noticed significant change in lessons after the professional development.

**Table 5: Evaluations with Findings on Teacher Knowledge but not Measurable Effects**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| JVA Consulting, LLC. (2006), **TQE/STEP**-CO | Gains in PRAXIS, PLACE; changes in attitudes as seen in participant teacher survey | Analysis of posttest results only | Treatment teachers only | Survey of teachers (N=53) showed 40 teachers agreed program prepared them to develop and utilize a variety of instructional approaches; post-treatment only passing scores from PRAXIS & PLACE exams show slight decline in both math and science from 2005 to 2006. |
| Lord, B. (2004), **CONNECT-Ed**-NJ | Attitudes and self-assessment as recorded in participant teacher survey | Analysis of posttest results survey results | Treatment teachers only | Survey of teachers indicated positive response that skills were gained. |
| Grip, R. S. (2006), **TaLL**-NJ | Attitudes and self-assessment as recorded in participant teacher survey, teacher end-of-course grades | Analysis of pre-/posttest results from middle school teachers | Treatment teachers only | Survey focused on response to PD activities over two years; from end-of-course grades 68.2% of participant teachers (N=22, small sample) received a "B" or better. |
| Howard, M. (2006) **Launch II**-FL | Participant teacher post-summer institute attitudes survey | Analysis of pre-/posttests survey results | Treatment teachers only | 75% of teachers report being more knowledgeable in math/science content after PD. |
| Wilsman, M. J. (2005), **NW Wisconsin** | Attitudes and self-assessment as recorded in participant teacher survey | Analysis of pre-/posttests survey results | Treatment teachers only | Teachers report they gained knowledge of science & math content, but small sample and low response rate. |
| Rusch, T. L., et al. (2006), **Comprehensive Database PD**-OH | Program-specific assessment of participant teachers | Analysis of pre-/posttests | Treatment teachers only | Some positive results on conceptual knowledge gain, but low response rates. |

**Teacher knowledge with no measurable effects.** The six studies reported in Table 5 were found not to include measurable effects of professional development on teacher knowledge. The three primary reasons were study design, weak or indirect measures of knowledge, or size of sample and response rate from teachers. None of these six studies included a control or matched comparison group of teachers, and it is not possible to determine if any reported knowledge gain can be attributed to the professional development. Several studies used indirect measures such as teacher surveys on what they had gained from the professional development. These kinds of general response items to professional development may give information at a general level but it is not specific information that can determine a direct relationship between the program and types of knowledge gained. The grades received by teachers in their courses are also indirect measures of knowledge, since grading is hard to compare across program and institution. Studies with small samples of teachers make it very difficult to generalize findings since the program effects may be specific only to this small group. This problem is accentuated if the response rate among participants is not high.

**Table 6: Instructional Practices Effects: Evaluations Reporting Measurable Effects of Professional Development on Teacher Instructional Practices**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| Howard, M. N. (2006), **Launch II**-FL | Changes in instructional methods as shown in participant teacher survey with 100-pt. scale of practice indicators | Analysis of pre-/posttests | Treatment teachers only | Decrease in traditional methods (from 59% to 25% use traditional methods); increase in standards-based methods (from 19% to 28%). |
| Brendefur, J., et al. (2006), **Developing Math Thinking**-ID | Correlation of gains in teacher knowledge inventory and instructional changes as observed in the classroom | Analysis of pre-/posttests | Treatment teachers only | Strong, statistically significant relationship between teachers' content knowledge gains and instructional practices (2-tail test, Pearson correlation coefficient=.327, significant at p <.05). |
| Niess, M. L. (2006), **High Desert**-OR | Changes in instruction as recorded in RTOP teacher observation | Content analysis of and scoring of classroom observations | Treatment teachers only - classroom observations | High average post scores on RTOP scales (N=32 observations): lesson design/implementation (13.728), content-propositional (15.548), content-procedural (12.758), communicative interactions (13.099), student/teacher relationships (16.365), and total (71.498). |
| Weaver, D. (2007), **OMLI**-OR | Changes in instruction as recorded in RMC Teacher Survey with class practices index | Analysis of pre-/posttests results | Treatment (School Leadership Team or SLT) teachers and comparison teachers | SLT teachers reported statistically significant decrease in use of traditional teaching practices across the two years the survey was conducted (2005-06 and 2006-07 (N=110, mean difference=6.382, significant at p<.05). |

## Instructional Practices

**Changes in Instructional Practice as Effects of Professional Development (4 studies).** The four studies shown in Table 6 that reported measurable effects on instructional practices had several common characteristics that give the reader confidence that the reported findings are accurate and valid. Three criteria were used to determine which study findings on instructional practices had measurable effects:

- Practice could be linked to teacher PD experienced;
- Measure of practice is close to the classroom (e.g., not close = teacher attitudes about practice); and
- Change in practice is measured for same teacher.

First, all four studies used data collection instruments that had been previously developed and tested for validity in measuring instructional practices in classrooms. Data from each of the instruments could be reported in scale measures that aggregate items to measured constructs. Two of the studies (Howard, 2006; Weaver, 2007) measured change in traditional teaching methods as compared to standards-based methods (as defined within the state). All four studies used pre-post measures for the treatment group and the data could be statistically analyzed using multi-item scales from the baseline point to a follow-up point after professional development. Each of the studies had sufficient numbers of same teachers in the sample at the pre- and post-test points who could be linked to the professional development treatment so that change in practices could be tested for statistical significance.

The Brendefur evaluation study of DMT professional development provided a correlation analysis of the relationship between gains in teacher knowledge and change in instructional practices. This method of relating change between evaluation measures is a good example of positive development in professional development evaluation.

**Table 7: Findings on Instructional Practices but not Measurable Effects**

| Evaluation Report, PD | Outcome Measure | Study Design | Treatment/ Comparison | Finding |
|---|---|---|---|---|
| Meta Associates. (2007), **NE Front Range**-CO | Participant teacher surveys, responses to open-ended questions | Content analysis and empirical coding of responses using NVivo | Treatment teachers only | 648 of 1405 open-ended responses reported changes in teacher attitudes toward teaching and student learning, application of pedagogies taught and modeled in the institute, and other teacher gains (including content knowledge). Of the 648 responses 58% dealt with changes in pedagogy, including teachers applying lessons learned from PD, instructional practices changing toward inquiry learning, changes in assessing students, use of student group work, and use of questioning. |
| Heath, et al. (2007), **Rocky Mountain MS**-CO | Changes in instruction as recorded in RTOP teacher observation | Content analysis of and scoring of classroom observations | Treatment teachers only | In 2006 observations of a small sample (N=6) of participant teachers, all scored between 66 to 96, showing at least that some evidence of reformed teaching practices was evident, and all showed improvement from 2005 observations. |
| Rosenblum, J. (2006), **MATHS**-ME | Changes in Instruction as indicated by Classroom Observation Protocol | Content analysis of and coding of classroom observations, pre-/post-treatment | Treatment teachers only | In 2004-05, nine classes with 534 total mins. observations accounted. In 2005-06, 27 classes with 1,658 total mins. accounted. A comparison between the two sets of observations showed an increase in range of instructional strategies used and an increase toward higher levels of cognitive thinking with two instructional types: teachers interacting with students and teacher lectures with discussion. |
| Grip R.S. (2006), **TaLL**-NJ | Participant teacher survey – open-ended responses to planned use of learning from course | Analysis of post-treatment survey responses | Treatment teachers only | 59.1% of the teachers that attended the math institute (N=22) responded that they will use/are using the math content element taught in their PD in their middle school classrooms. Similarly, 60.0% of teachers who participated in the geometry institute (N=10) said they will apply or are applying what they learned in their own classrooms. |
| McKnight, K. (2006), **BreakThrough Math**-OH | Teacher responses to post-treatment participant teacher survey | Analysis of post-treatment survey | Treatment teachers only | 72% of teachers that responded to the survey (N=97) said that they have made changes in their teaching that they attribute to their course with BreakThrough Math. |

**Instructional practices with no measurable effects:** Unlike the four studies described in Table 6, above, the five studies reported in Table 7 did not produce measurable effects, even though positive results are described. One key reason is the lack of design rigor: across all five studies the measures of practice were applied only to the treatment teachers. It is unknown and thus possible that comparison teachers would also produce similar positive results without any link to the teacher PD that was experienced by the treatment teachers. Another is the small sample size used: both the Rocky Mountain Middle School MSP and TaLL programs have sample sizes under 30 which limit the generalizability of the findings. Moreover, none of the studies provided statistical analysis to discount for other potential factors or discount the possibility of chance that could contribute to the positive results. Lastly, in three of the five studies teachers self-reported on either a change in attitude about instructional practice or in their intent to apply what they have learned. Although a change in attitude or intent is a step toward changed instructional practice, it is still a distal indicator from actual measurements of changed instructional practice as experienced by the teachers and students.

**Professional Development Designs with Measurable Effects: *What have we learned about how these programs operate?***

We can compare and analyze the characteristics of eight professional development programs that were identified through the evaluation studies as having measurable effects on student outcomes (7 program designs) or instructional practices (4 program designs).

In Table 8, we identify several characteristics of how the professional development was organized and delivered to produce effects on teaching and learning. We are focusing our analysis on the professional development that was provided in the studies documenting measurable effects. Thus, we want to focus in on the characteristics of the program designs and activities that are likely to produce effects. What do we see in common across the programs that made them effective?

A common feature of the eight program designs was relatively high amount of time for each teacher in professional development (varying from 45 hours to over 300 hours in annual time, or nine to 12.5 graduate credit hours per year). Most programs were designed for over 100 hours of teacher activities in professional development. All but two programs included a two-week intensive summer institute design for increasing content and pedagogical knowledge of teachers in math or science.

All eight of the programs targeted teachers in elementary grades or elementary and middle grades math or science. The designs also included significant activities during the school year that focused on coaching and mentoring through master teachers, use of the lesson study method for learning among teacher colleagues, or further content or pedagogy instruction during periodic training sessions. One Oregon program required two full days of math content and pedagogy instruction each Friday and Saturday during the school year. Thus, the program designs did not focus only on teachers learning more math or science content, or *primarily* on development of content knowledge, but rather they were described as emphasizing knowledge of how to teach content to the students.

We also note in the descriptions of these eight professional development designs a commitment to linking the teacher development to the curriculum and organization of their assigned schools. Research on effectiveness of professional development shows programs that result in change in teaching practices are characterized by coherence to the curriculum objectives of the schools as well as collective participation among teachers in the learning process. The eight programs with measurable effects all have design elements and activities that bring teachers together to continue their education and build their knowledge—through lesson study development, observing and reflecting with colleagues, or designing the treatment activities with grade level teacher teams. In each of these programs the schools were a strong partner in building and implementing the PD.

**Table 8: Characteristics of Teacher Professional Development Designs that show Measurable Effects – Student Achievement or Instructional Practices**

| Evaluation Report, PD | Method of PD | Content of PD, Learning goals | Providers & Activities | Duration and Contact Hours, Treatment Group Size |
|---|---|---|---|---|
| Howard M. (2006), **Launch II**-FL | Focus on elementary and middle schools in the same feeder pattern, Gr. 3-8 teachers; summer institutes and follow-up activities through next school year | Content knowledge and instructional skills for math and science using "space" as an integrating theme; participating schools commit to building and sustaining a leadership cadre for mathematics and science | Providers: Florida Gulf Coast Univ., Florida Space Research Institute and NE Florida Educational Consortium.<br>Activities: Two-week summer institutes, in-school activities engaged participants, district-based workshops, online learning modules, and one-on-one mentoring from staff | 1 year, 60 hours,<br>150 Gr. 3-8 teachers in math and science/year |
| Brendefur, J., et al. (2006), **Developing Math Thinking**-ID | Elementary teacher-administrator teams focus on math content and assessment; University summer institutes plus in-school weekly activities | Increase elementary teachers' capacity to teach for understanding, focusing on conceptual understandings and misunderstandings behind arithmetic calculations; uses cognitively guided instruction (CGI) framework | Provider: Boise State University<br>Activities: One-week summer institute with university math faculty with teacher- administrator teams; weekly meetings with university faculty that include development and use of multiple student assessments and analyses of the data | 3 years, 135 hours, 56 teachers and four administrators |
| Niess, M. L. (2005), **High Desert**-OR | Math content and pedagogy for Gr. 3-8 teachers in full-year program, using off-site (university) and on-site (school-based) activities | Prepare teachers with rigorous math content, pedagogical content knowledge, and collaborative techniques; target those teaching minority and disadvantaged students not meeting state standards in mathematics | Provider: Oregon State University<br>Activities: Two-week summer institute with higher education partnered with districts; three terms of math courses & pedagogy (Fri-Sat, 8 hrs./week); in-school modeling and observation with master teachers | 1 year, 304 hours, 25 Gr. K-8 math teachers |
| Weaver, D. (2007), **OMLI**-OR | Develop math leader teams in schools through summer institutes and school-based activities | Goal is development of a cadre of math teachers as school- and district-based intellectual leaders and master teachers through intensive summer institutes and follow-up academic year activities; train two teachers and one administrator (per school) to work as change agents for math instruction within schools | Providers: OR State Univ., Portland State University, Teacher Development Group.<br>Activities: Institutes combine rigorous math content coursework with leadership development in seminars; academic year activities facilitate professional learning communities K-12; teacher content emphasizes student discourse, and teachers learn model pedagogical techniques | 3 years, 27 graduate course credit hours or 360 hours.<br>280 elementary, middle, high teachers and administrators in math |

**Table 8 – continued**

| Evaluation Report, PD | Method of PD | Content of PD, Learning goals | Providers & Activities | Duration and Contact Hours, Treatment Group Size |
|---|---|---|---|---|
| Meta Associates. (2007), **NE Front Range**-CO | Summer institutes plus lesson study with teacher teams | Develop content knowledge of target group of teachers in specific content areas: analysis & probability, geometry, earth/space science, forces & motion, and life science | Providers: Colorado School of Mines, Univ. of Colorado, Boulder & Denver, University of Denver, Univ. of N. Colo., Denver Museum of Nature and Science<br>Activities: Summer institutes with five districts; four follow-up institutes with same faculty and content specialists; teachers participate in lesson study, viewing and critiquing a lesson taught by a colleague to different groups of students | 1 year, 120 hours<br>70-150 Gr. 6-8 teachers in math and science/year |
| Evansville-Vanderburgh Schools. (2006), **iCATS**-IN | Monthly workshops for teachers by grade level; focus on Gr. 3-5 teachers in schools with lower than expected math scores in Gr. 6 | Facilitated interventions in schools' grade level teams that provide training and support for teachers to use hands-on approach to teaching math and use manipulatives and technology | Provider: District Curriculum and Technology Specialists<br>Activities: Workshops focus on teaching math concepts, modeling the methods to be used in the classroom with grade level teams. Teachers construct and teach lessons based on State Standards and specific mathematical concepts. Lessons all posted on the web for reflection and discussion | 1 year, 45 hours<br>50 Gr. 3-5 teachers/year |
| Grip R. S. (2006), **TaLL**-NJ | Graduate courses in math content and methods for middle grade teachers | Graduate credits in mathematics and teaching methods toward MA. Emphasis on NJ state standards; teachers get math specialization and highly qualified status | Provider: The College of New Jersey, Schools of Education and Science<br>Activities: Courses offered to middle grade teachers in high risk schools, with non-participating teachers serving as comparison for evaluation | 3 years, 37.5 graduate credit hours, 100-200 Gr. K-8 math teachers/year |
| Hansen, J. B. (2006), **Willamette Valley**-OR | Summer institutes plus lesson study with teacher teams | Science teachers design a lesson for a two week unit in the classroom based on content in summer workshop and observe/ work with treatment colleagues when they introduce new content. | Providers: Oregon State University and the Teaching Research Institute of Western Oregon University<br>Activities: Summer institute and three release days during school year to work with other teachers in program | 1 year, 9 graduate credit hours<br>15 Gr. 5-8 math teachers/year |

## Instruments Used in Evaluations

As shown in Table 9, nearly all of the studies that had measurable effects in student achievement outcomes utilized readily available statewide standardized assessments. The two exceptions – Willamette Valley and OMLI – differed due to their specific content area. Due to limitations in the state assessment aligning with the watershed curriculum, Willamette Valley contracted a known assessment developer to provide them with their assessment. On the other hand, OMLI was interested in moving student discourse toward inquiry and elevating them to higher cognitive levels, which the current state standardized assessment does not measure. The Student Discourse Observation Protocol was thus developed based upon prior research, including those stemming from work with the National Science Foundation's Local Systemic Change Initiative. Key to continued use of statewide standardized assessments is alignment with the content area of the professional development and with the classroom curriculum.

Program-specific assessments dominate the instruments used in capturing teacher knowledge and change in instructional practice in evaluations with measurable effects. Nevertheless, there is a small but emerging development in the use of widely tested instruments to measure teacher content knowledge and instructional practices, such as DTAMS, CKT-M, PRAXIS, RTOP, and the Surveys of Enacted Curriculum.

**Table 9: Instruments for Measuring Outcomes of Professional Development**

| Evaluation Report, PD | Student Achievement Outcome Instrument | Instructional Practice Outcome Instrument | Teacher Knowledge Outcome Instrument |
|---|---|---|---|
| META Associates. (2007), **NE Front Range**-CO | **Standardized State Assessment:** CSAP | **Program-Specific Assessment:** Participant Teacher Surveys | **Program-Specific Assessment:** Content-Based Teacher Assessment |
| Heath, et al. (2007), **Rocky Mountain MS**-CO | | | **Program-Specific Assessment:** Teacher Content Inventory |
| Howard, M. (2006), **Launch II**-FL | | **Program-Specific Assessment:** Participant Teacher Surveys, scales for practices | |
| Schmidt, D. L. (2006), **Launch II**-FL | **Standardized State Assessment:** FCAT | | **Program-Specific Assessment**: Teacher Content Knowledge Assessment |
| Brendefur, J., et al. (2006), **Developing Math Thinking**-ID | | **Program-Specific Assessment:** Teacher Knowledge Inventory, Classroom Observation | **Program-Specific Assessment**: Number Knowledge Inventory |
| Evansville-Vanderburgh Schools. (2006), **iCATS**-IN | **Standardized State Assessment:** Indiana Curriculum Framework Assessment | | |
| Chen, A. (2006), **Edutron**-MA | | | **Standardized, State Assessment:** Massachusetts Tests for Educator Licensure (MTEL) **& Program-Specific Assessment**: course assessments |
| Lesley University. (2006), **Coalition for Higher Math Standards**-MA | | | **Program-Specific Assessment:** Course Assessments |

**Table 9 – continued**

| Evaluation Report, PD | Student Achievement Outcome Instrument | Instructional Practice Outcome Instrument | Teacher Knowledge Outcome Instrument |
|---|---|---|---|
| Grip, R. S. (2006), **TaLL**-NJ | **Standardized, District-wide Assessment:** Terra Nova | **Program-Specific Assessment:** Participant Teacher Surveys – Planned use of learning from course | **Program-Specific Assessment:** End-of-Course Grades |
| Hankerson, L. (2006), **Nash Rocky Mount**-NC | | | **Standardized Assessment:** PRAXIS |
| Hansen, J. B. (2006), **Willamette Valley**-OR | **Program-Specific Assessment:** Student Science Achievement Level Test (developed with Northwest Evaluation Assoc.) | | **Program-Specific Assessment:** Faculty-Developed Assessment; Inquiry-Based Teaching Rubric (based on standards developed by Chief State Science Supervisors) |
| Weaver, D. (2007) **OMLI**-OR | **Program-Specific, Tested Assessment:** Student Discourse Observation Protocol (developed by RMC Corporation) | **Program-Specific Assessment:** RMC Teacher Surveys, Class Practice Index | **Standardized, Tested Assessment:** Content Knowledge Assessment (CKT-M) |
| Niess, M. L. (2005) **High Desert**-OR | **Standardized State Assessment:** Knowledge & Skills Test | **Standardized Assessment:** RTOP, Surveys of Enacted Curriculum | **Standardized, Tested Assessment:** DTAMS, PRAXIS |

# Conclusions and Recommendations

The study was developed in response to several key issues for state education leaders regarding questions about teacher professional development in math and science, and how states can take a stronger leader role in improving program quality. From findings across the evaluation studies, CCSSO has identified several key areas for state action:

1) **One-third of evaluation studies reported measurable effects of teacher professional development.** Seven of the evaluation studies of teacher professional development reviewed by CCSSO reported measurable effects of the teacher development activities on subsequent student outcomes. A total of 10 of the studies reported measurable effects on increasing teacher content knowledge, and four studies reported measurable effects on instructional practices of teachers.

2) **Significant effects in programs designed with content-focused PD plus sufficient time plus in-school component.** The cross-program review of studies showed significant effects of professional development programs for teachers of math and science when the programs include a focus on content knowledge in the math and science subject areas plus training and follow-up pedagogical content knowledge. The total time in professional development for the studies with significant effects was 50 hours or more. Most program designs are intended to provide about one year of teacher involvement in a specific set of activities. Then, many programs are designed to begin with a new cohort of teachers. A key question is whether teachers have sufficient time to fully implement what may have been learned, and whether follow-up activities with teachers are sufficient to have an impact on improving teaching. The CCSSO analysis showed that programs with findings showing significant effects on teaching and learning had more time focused on teacher development in math or science content, such as summer institutes, but they also featured major follow-up efforts with teachers in the classroom or school so that teaching practices learned could be reinforced and improved after the teachers had begun to try them with students.

3) **Important to plan purposeful evaluations that yield measurable effects.** This study began with the core assumption that the field did not have good evidence that evaluation of program effects were a key element of professional development in most funded programs. Moreover, federally funded initiatives called for evaluations and reporting of findings to the federal offices, but there is a dearth of good evidence of how well-designed and executed these evaluations are. The evidence from the CCSSO review of evaluation studies shows that one-third of the programs reviewed did have well-developed evaluations that produced findings with measurable effects on student achievement or change in instructional practices. When the programs were reviewed for quality elements, based on research, a majority of the programs were highly rated. Strong evidence of effects of these programs on student learning or on changing teaching was gained from only one-third of the programs. Our analysis of evaluation findings emphasized scientific study design, and these kinds of designs could have been implemented across more programs.

4) **Build valid, tested instruments into the evaluation design.** A third issue for the study was how evaluations of professional development are designed in relation to the purposes for the evaluations. A majority of programs planned to use student achievement data to evaluate effects of professional development. About half the programs that reported results did include multiple years of achievement data that could be analyzed to determine change. However, in many cases the link could not be made between teachers that received professional development and the students that they subsequently taught (i.e., pre-post measurement). Finally, some of the evaluation studies that were reviewed emphasize methods that require extensive evaluator staff time, such as classroom observations. These methods provide more in-depth qualitative data and

they may be useful on a small scale at initial stages of implementation of a program, but they become very expensive as programs are expanded.

5) **Weigh carefully teacher-based vs. school-based design for developing knowledge and skills.** Many of the professional development designs for teachers reviewed in the study were based on selection through teacher-based, voluntary models, even though specific districts and schools were targeted. Exceptions to the pattern were the school-based designs in Launch II, OMLI, DMT, and iCATS. The use of teacher-based professional development makes important follow-up activities harder to schedule and implement, and alignment to school curriculum more difficult to accomplish. For evaluation, the use of student assessment scores and tracking change over time appear to be facilitated with use of the school-based model for professional development. Additionally, less time and resources needs to be expended in tracking the time teachers spend in professional development since participation is largely defined by the school. The question of PD quality or fidelity of implementation still is important for evaluation.

6) **Include outcomes measures in allocation of evaluation resources.** Larger professional development programs that also span multiple years of funding support, such as those supported through NSF MSP projects typically have multiple methods and stages of evaluation, including quality of implementation, teacher knowledge gains, change in instruction, and improvement of student achievement. The larger programs also can afford on-site observation and formative evaluation reports to assist in mid-course corrections of program designs. Some projects plan to conduct research studies with the data from the program. Smaller professional development projects typically cannot afford ambitious, multi-stage evaluations or research. Allocation of more funds to evaluation would mean fewer participating teachers or fewer resources for the program implementation. In the cross-state program review we observe that smaller programs typically had to choose a few measures and methods of evaluation. We noted that several evaluation reports showing measurable effects devoted less attention and resources to evaluating implementation and on-site observation studies, and focused more evaluation effort with measuring outcomes from the professional development.

7) **Plan for use of data systems and experimental designs.** Evaluations that will measure effects over time require access to data collection instruments or data systems, and advance planning with school officials. About one-fourth of programs evaluations did include comparison of a treatment group with a control group of teachers. Our analysis suggests that many professional development programs could use a scientific design without great extra costs. State managers should consider evaluation designs that can be completed by linking data from state student assessments or local assessments with data on professional development for teachers. Integrating data systems will remove the need for buying or creating new evaluation tools for each program to be evaluated. To build a matched comparison group or control group of teachers into the evaluation design requires significant cooperation and buy-in to the benefits of evaluation from school administrators and district leaders. Often, advance permission is needed from teachers to collect data on instructional practices that can be compared between treatment and control groups or to link measures of teacher knowledge to subsequent measures of practices or student achievement. Thus the role of the evaluation has to be carefully explained as providing larger benefits for school systems due to evidence that will be gained, and how it will ensure better decision-making in the future.

8) **Link teacher knowledge gains to change in classroom practices.** One type of evaluation finding identified in this review of studies showing promise for further use and expansion to other PD studies was measurement of change in teaching practices in the classroom. Four of the studies implemented well-tested instruments for comparing classroom practices across samples of

teachers and classrooms. The evaluators accessed instruments that could be used more broadly, such as RTOP and SEC. With advance planning, teachers and classrooms can be selected so that change in practices can be measured at the baseline point when teacher development begins and after implementation of activities and a period of implementation has been experienced. Gaining cooperation from a sample of comparable teachers to measure their change over the same period is very important for making attribution to the PD initiative. Two issues need to be considered: a) allowing sufficient time to measure effects on classrooms; b) determining the goal for change in practices. One option for measuring change is determining change in relation to the objectives of the professional development (e.g., improve teaching of geometric concepts in middle grades), and a second option is measuring change in relation to content standards or curriculum framework for the subject and grade (i.e., professional development improved coherence of curriculum).

9) **Consider timely use of findings in program decisions by key decision-makers.** The program descriptions and evaluation reports that were reviewed by CCSSO for this study did not focus on dissemination plans or how the results from the evaluation studies were intended to be used. However, a CCSSO survey conducted in June-July 2006 suggests most of the program proposals **did** describe plans for sharing their evaluation data with district and state officials (22 out of 25) and grant donors (20 out of 25). Fifteen of the 25 programs intend to share evaluation results with the participant teachers and 18 of 25 said they would share with school administrators. On rare occasions would dissemination include parents, students or school board members as stakeholders. In addition, most of the programs intended to disseminate their evaluation results using traditional avenues: white papers (19), formal presentations (17), as well as informal presentations (17) and conversations (18).

   We do see evidence in the study reports of internal evaluation results and formative feedback to program managers primarily related to how activities were implemented, and these kinds of results can be extremely valuable. However, with the recent new attention to scientific designs to provide measurement of impact of professional development on learning, we would like to see greater focus on how results from evaluations will be provided to decision-makers at specific points of time in the course of a project, and not long after the program activities have concluded. This is particularly important if a specific model is being considered for replication or expansion to other districts, schools, or additional teacher groups.

10) **Program/states should appraise the value of partnerships for evaluation.** From analyzing the evaluation findings across programs CCSSO has noted that the partnerships between higher education institutions and school districts have generally not added to the capacity for evaluation of professional development. Evaluation specialists and consultants from non-higher education organizations have led most of the evaluation designs, data analysis, and reporting. From the viewpoint of higher education institutions, a key partner for obtaining data and measures for evaluation are local school district decision-makers and state education agency officials. Access to data records for teachers and students is often a key step in implementing evaluation designs that provide tracking of professional development effects over time.

# Evaluation Study Reports Reviewed

**BreakThrough Math**

McKnight, K. (2006, June). *Evaluation summary report for teacher follow-up survey*. Ohio Department of Education BreakThrough Math project. Arlington, VA: LessonLab Research Institute, Pearson Achievement Solutions.

**Comprehensive, Data-Based PD**

Rusch, T. L., Farrell, A. M., & Reed, M. K. (2006, May 11). *Comprehensive, data-based professional development. Final report, narrative & data. Report to the Ohio Department of Education Middle Grade Mathematics Professional Development.* Dayton, OH: Wright State University.

**CONNECT-ED**

Lord, B. (2004, October). *CONNECT-ED pilot project*. Final evaluation report. Report to Rider University SELECT, Bristol-Myers Squibb Center for Science Teaching and Learning. Newton, MA: Center for Leadership and Learning Communities, Education Development Center.

**DMT**

Brendefur, J., Espinoza, I., & Pfiester, J. (2005). *Developing mathematical thinking. Project Micron Foundation grant and mathematics and science partnership grant. Technical report – year 1, July 1, 2004-June 30, 2005.* Boise, ID: Boise State University.

Brendefur, J., Espinoza, I., & Pfiester, J. (2006). *Developing mathematical thinking. Project Micron Foundation grant and mathematics and science partnership grant*. *Technical report – year 2, July 1, 2004-June 30, 2006.* Boise, ID: Boise State University.

**EduTron**

Chen, A. (2006). *Intensive immersion institute in mathematics for grades 4-8 teachers.* Report to the U.S. Department of Education, Mathematics and Science Partnerships. Winchester, MA: EduTron Corporation.

Chen, A. (2006, Fall). *Annual report addendum*. Report to Massachusetts Department of Education Massachusetts Math and Science Partnerships Program. Winchester, MA: EduTron Corporation.

**High Desert**

Niess, M. L. (2005). *Oregon ESEA Title IIB MSP: Central Oregon consortium.* Report to the U.S. Department of Education, Mathematics and Science Partnerships. Corvallis, OR: Department of Science & Mathematics Education, Oregon State University.

**iCATS: Beyond the Textbook**

Evansville-Vanderburgh School Corporation Integrating Curriculum and Technology Specialists. (2004, May). *Ed tech grant posttest data*. (Year 1). Unpublished manuscript. Evansville, IN: Author.

Evansville-Vanderburgh School Corporation Integrating Curriculum and Technology Specialists. (n.d.). *Ed tech grant posttest data* (Year 2). Unpublished manuscript. Evansville, IN: Author.

Evansville-Vanderburgh School Corporation Integrating Curriculum and Technology Specialists. (2006, July 10). *Ed tech grant posttest data*. (Year 3). Unpublished manuscript. Evansville, IN: Author.

**LAUNCH II**
Howard, M. (2006, September). *Project LAUNCH II external evaluation report.* Unpublished manuscript.

Schmidt, D. L. (2006). *Project LAUNCH II*. Report to the U.S. Department of Education, Mathematics and Science Partnerships. Fort Myers, FL: Whitaker Center at Florida Gulf Coast University.

**Lesley/MassInsight**
*Coalition for higher standards math partnership program.* (2006). Report to the U.S. Department of Education, Mathematics and Science Partnerships. Cambridge, MA: Lesley University.

Meyers, B. (2006, January 19). *Advanced math coaching in urban school districts.* Report to the Massachusetts Board of Higher Education, Improving Teacher Quality Higher Education Partnership Grant Program. Unpublished manuscript.

**M4**
Colonis, M., & Rodriguez, E. (n.d.). *M4 – making math more meaningful mathematics and science partnership grant k-9 mathematics initiatives.* Evaluation report I, 2005-06 academic year. Report to the Superintendent, Crawford Community School Corporation and Grant Coordinator, Crawford Community School Corporation. Unpublished manuscript.

**MATHS**
Rosenblum, J. (2005, September). *Mathematics: Access and teaching in high schools – MATHS, 2004-05 evaluation report.* Augusta, ME: Maine Mathematics and Science Alliance.

Rosenblum, J. (2006, August). *Mathematics: Access and teaching in high schools – MATHS, 2004-06 evaluation report.* Augusta, ME: Maine Mathematics and Science Alliance.

**MS*TEAMS**
Fiene, J. (2006). Tiered mentoring project year 2 report. Report to the Green River Regional Educational Cooperative. Unpublished manuscript.

**Nash-Rocky Mount**
Hankerson, L. (2006, October 4). *Pursuing excellence in middle school math & science*. Project Narrative. Report to the U.S. Department of Education, Mathematics and Science Partnerships. Nashville, NC: Nash-Rocky Mount Schools.

Perry, M. (2005, July 28). *Pursuing excellence in middle school math & science.* Project Narrative. Report to the U.S. Department of Education, Mathematics and Science Partnerships. Nashville, NC: Nash-Rocky Mount Schools.

Perry, M. & White, C. T. (2006, October 4). *Pursuing excellence in middle school math & science.* Project Profile. Report to the U.S. Department of Education, Mathematics and Science Partnerships. Nashville, NC: Nash-Rocky Mount Schools.

White, C. T. (2006, October 30). *Percent proficient on Spring 2006 EOG reading and math tests & alternate assessments.* (Data table). Unpublished manuscript.

White, C. T. (2007, January 2). *MSP impact and effectiveness of the mathematics and science coaching program.* Unpublished manuscript.

**Northeast Front Range**

META Associates. (2004, November). *Northeast Front Range math/science partnership (MSP) to increase teacher competence in content. Year 1 evaluation report.* Golden, CO: Author.

META Associates. (2006, March). *Northeast Front Range math/science partnership (MSP) to increase teacher competence in content.* Year 2 evaluation report: January 1, 2005–December 31, 2005. Golden, CO: Author.

META Associates. (2007, March). *Northeast Front Range math/science partnership (MSP) to increase teacher competence in content.* Final evaluation report, January 1, 2004–December 31, 2006. Golden, CO: Author.

**NW Wisconsin**

Wilsman, M. J. (2005, December). *Summary evaluation report for the northwest Wisconsin mathematics and science partnership*. Unpublished report, University of Wisconsin-Madison.

Wilsman, M. J. (n.d.). *Northwest Wisconsin mathematics and science partnership. Evaluation report on the July 2004 course.* Unpublished report, University of Wisconsin-Madison.

**OMLI**

Weaver, D. (2006, May). *Oregon mathematics leadership institute.* Spring 2006 evaluation report. Report to the Oregon State University, Department of Mathematics. Portland, OR: RMC Research Corporation.

Weaver, D. (2007, April). *Oregon mathematics leadership institute.* Spring 2007 evaluation report. Report to the Oregon State University, Department of Mathematics. Portland, OR: RMC Research Corporation.

**R & R**

Greenseid, L., & Lawrenz, F. (2005, June 9). *WASDI retention and renewal program evaluation: Second year executive summary*. Unpublished report, University of Minnesota, Minneapolis.

**Rocky Mountain Middle School**

Heath, B. P. (2006, May 18). *Rocky Mountain middle school mathematics and science partnership.* Annual evaluation report: Year 2 final. Wilmington, NC: East Main Educational Consulting, LLC.

Heath, B. P., Lakshamanan, A., Perlmutter, A., Davis, L., Whaley, B., Walthers, B. J., et al. (2007, May 16). *Rocky Mountain Middle School mathematics and science partnership. Section 4 evaluation annual report year 3: 2006-07*. Wilmington, NC: East Main Educational Consulting, LLC.

**Science Coalition Courses and Science Lead Teacher**

Noble, A. J., Boudreau, C., & Smith, A. (2000, October). *Local systemic change initiative evaluation report.* Publication No. T00.014.1. Newark, DE: Delaware Education Research & Development Center, University of Delaware.

Wood, R. E., & Collette, J. W. (n.d.). *The Delaware LSCI: A systemic partnership to improve the teaching and learning of science for all children.* Final report to the National Science Foundation. Unpublished manuscript.

**South Carolina Coaching Initiative**

Michel, Y. (2007). *Report of teacher surveys in coached schools 2003*. Report to the South Carolina State Department of Education Math and Science Unit. Unpublished manuscript.

**TaLL**

Grip, R. S. (2005, October). *Year-end evaluation of the teacher as leaders and learners program.* Mount Laurel, NJ: Statistical Forecasting, LLC.

Grip, R. S. (2006, October). *Year-end evaluation of the teacher as leaders and learners program.* Mount Laurel, NJ: Statistical Forecasting, LLC.

**TQE/STEP**

JVA Consulting, LLC. (2005, December). *Teacher quality enhancement grant.* Annual evaluation progress report: Year 1. Denver, CO: Author.

JVA Consulting, LLC. (2006, October). *Teacher quality enhancement grant*. Annual evaluation progress report: Year 2. Denver, CO: Author.

**Willamette Valley**

Hansen, J. B. (2005, September). *Evaluation of the Willamette Valley Watershed Partnership project – 2004-05.* Year one external evaluation report. Monmouth, OR: The Teacher Research Institute's Evaluation Research Group Office, Western Oregon University.

Hansen, J. B. (2006, October 17). *Evaluation of the Willamette Valley Watershed Partnership project – 2005-06.* Year two external evaluation report. Monmouth, OR: The Teacher Research Institute's Evaluation Research Group Office, Western Oregon University.

Willamette Educational Service District. (2006, October 17). *Willamette Valley Watershed Partnership project.* 2005-06 final report, part 2: Narrative responses. Unpublished manuscript.

# References


Banilower, E. R., Boyd, S. E., Pasley, J. D., & Weiss, I. R. (2006, February). *Lessons from a decade of mathematics and science reform: A capstone report for the Local Systemic Change through Teacher Enhancement Initiative.* (Research Monograph-Prepublication copy). Chapel Hill, NC: Horizon Research, Inc. Retrieved March 21, 2006, from http://www.pdmathsci.net/reports/capstone.pdf

Birman, B. F., & Porter, A. C. (2002). Evaluating the effectiveness of education funding streams. *Peabody Journal of Education, 77*(4), 59–85.

Blank, R. K., de las Alas, N., & Smith, C. (2007, February). *Analysis of the quality of professional development programs for mathematics and science teachers: Findings from a cross-state study*. Washington, DC: Council of Chief State School Officers. Retrieved February 29, 2008, http://www.ccsso.org/content/pdfs/year%202%20new%20final%20NSF%20Impde%20Fall%2006%20%20Report%20-032307.pdf

Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement* (CPRE Research Report No. RR-043). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Cohen, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform works.* New Haven, CT: Yale University Press.

Corcoran, T., & Foley, E. (2003). The promise and challenge of evaluating systemic reform in an urban district. *Research Perspectives on School Reform: Lessons form the Annenberg Challenge.* Providence, RI: Annenberg Institute at Brown University.

Council of Chief State School Officers (2006). *Rubric for professional development program review.* Retrieved March 05, 2007, from http://www.ccsso.org/content/PDFs/Revised%20Rubric%20%2D%20rev022406.xls

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81–112.

Frechtling, J. (2001). What evaluation tells us about professional development programs in math and science. In C. R. Nesbit, J. Wallace, D. K. Pugalee, A. Miller, & DiBiase, W. J. (Eds.), *Developing teacher leaders: Professional development in science and mathematics*. Columbus, OH: Education Resources Information Center.

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

Garet, M. S., Birman, B. F., Porter, A. C., Desimone, L., Herman, R. & Yoon, K. S. (1999). *Designing effective professional development: Lessons from the Eisenhower program and technical appendices* (Report No. ED/OUS99-3). Washington, DC: American Institutes for Research. (ERIC Document Reproduction Service No. ED442634)

Grant, S. G.. Peterson, P. L., & Shojgreen-Downer, A. (1996, Summer). Learning to teach mathematics in the context of systemic reform. *American Educational Research Journal, 33*(2), 509-41.

Hiebert, J. (1999, January). Relationships between research and the NCTM standards. *Journal for Research in Mathematics Education, 30*(1), 3-19.

Horizon Research, Inc. (2000, November 30). *Inside the classroom observation and analytic protocol.* (COP). Chapel Hill, NC: Author. Retrieved March 16, 2007, from http://www.horizon-research.com/instruments/clas/cop.pdf

Kennedy, M. M. (1998, April). *Form and substance in inservice teacher education*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Lieberman, A. (1996, November). Creating intentional learning communities. *Educational Leadership, 54*(3), 51-55.

Loucks-Horsley, S., Hewson, P., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics.* Thousand Oaks, CA: Corwin Press.

Weiss, I. R., Banilower, E. R., McMahon, K. C., & Smith, P. S. (2001). *Report of the 2000 national survey of science and mathematics education.* Chapel Hill, NC: Horizon Research, Inc.

# Appendix: Contacts for Reports

| Program Name | Local Coordinator(s) | Local Evaluator(s) |
|---|---|---|
| **BreakThrough Mathematics** | ***Kathy McKnight***<br>T: 703-389-5401<br>E:<br>Kathy.McKnight@pearsonachievement.com | ***Kathy McKnight***<br>T: 703-389-5401<br>E:<br>Kathy.McKnight@pearsonachievement.com |
| **Comprehensive, Data-Based PD** | ***Ann Farrell***<br>T: 937-775-2193<br>E: ann.farrell@wright.edu | ***Tracy Rusch***<br>T: 937-775-3275<br>E: tracy.rusch@wright.edu<br><br>***Michelle Reed***<br>T: 937-775-2785<br>E : michelle.reed@wright.edu |
| **CONNECT-ED** | ***Kathleen M. Browne***<br>T: 609-895-5408<br>E: browne@rider.edu<br>W: http://www.rider.edu/2559_5881.htm | ***Brian Lord***<br>T: 617-618-2547<br>E: blord@edc.org |
| **DMT** | ***Jonathan Brendefur***<br>T: 208-426-2468<br>E: jbrendef@boisestate.edu | ***Jonathan Brendefur***<br>T: 208-426-2468<br>E: jbrendef@boisestate.edu |
| **EduTron** | ***Andrew Chen***<br>T: 781-729-8696<br>E: schen@edutron.com | ***Andrew Chen***<br>T: 781-729-8696<br>E: schen@edutron.com |
| **High Desert** | ***Maggie Niess***<br>T: 541-737-1818<br>E: niessm@onid.oregonstate.edu | ***Edith Gummer***<br>T: 503-275-9171<br>E: gummere@nwrel.org |
| **iCATS: Beyond the Textbook** | **Mike Russ**<br>T: 812-435-8546<br>E: mkruss@evsc.k12.in.us | ***Amy McBride Martin***<br>T: 812-488-2381<br>E: am85@evansville.edu |
| **Launch II** | ***Diane L. Schmidt***<br>T: 239-590-7741<br>E : dschmidt@fgcu.edu<br><br>***Donna Price Henry***<br>T: 239-590-7156<br>E: dhenry@fgcu.edu | ***Michael N. Howard***<br>T: 336-286-2683<br>E: michaelnhoward@earthlink.net |
| **Lesley/MassInsight** | ***Ann Collins***<br>T: 617-349-8853<br>E: acollin8@lesley.edu | ***Bud Meyers***<br>T: 802-656-3282<br>E: bud.meyers@uvm.edu |
| **M4** | ***Phyllis Boswell***<br>T: 765-362-2342<br>E: pboswell@admin.cville.k12.in.us | ***S. Betsy Berry***<br>T : 765-414-0216<br>E : bestberry@perdue.edu |
| **MATHS** | ***Cheryl Rose***<br>T: 207-287-5970<br>E: crose@mmsa.org<br><br>***Francis Eberle***<br>E: francis.eberle@gmail.com | ***Jill Rosenblum***<br>T: 207-287-6644<br>E: jrosenblum@mmsa.org |
| **MS*TEAMS** | ***Liz Storey***<br>T: 270-745-2451<br>E: liz.storey@grrec.ky.gov | ***Jeanne Fiene***<br>T: 270-745-4890<br>E: jeanne.fiene@wku.edu |
| **Nash-Rocky Mount** | ***Carol Turner White***<br>T: 252-459-5329<br>E: ctwhite@nrms.k12.nc.us<br><br>***Michael Perry***<br>T: 252-459-5255<br>E: mdperry@nrms.k12.nc.us | ***Carol Turner White***<br>T: 252-459-5329<br>E: ctwhite@nrms.k12.nc.us<br><br>***Michael Perry***<br>T: 252-459-5255<br>E: mdperry@nrms.k12.nc.us |
| **Northeast Front Range** | ***Linda J. Morris***<br>T : 303-982-6955<br>E : ljmorris@jeffco.k12.co.us<br><br>***Marsha Barber***<br>T: 303-989-3255<br>E: marshabarber@mywdo.com | ***Cari Swenson***<br>T: 303-526-7480<br>E: cari@meta1.us<br><br>***Susan Duron***<br>T: 303-526-7480<br>E: duron1@aol.com |

**Appendix: Contacts for Reports – continued**

| Program Name | Local Coordinator(s) | Local Evaluator(s) |
|---|---|---|
| **NW Wisconsin** | ***Julie C. Stafford***<br>T: 715-723-1181<br>E: jstafford@wasdinet.org | ***Margaret J. Wilsman***<br>T: 608-263-4541<br>E: mwilsman@wasdinet.org |
| **OMLI** | ***Thomas P. Dick***<br>T: 541-737-1570<br>E: tpdick@math.orst.edu | ***Dave Weaver***<br>T: 503-223-8248<br>E: Dave_Weaver@rmccorp.com |
| **R & R** | ***Julie C. Stafford***<br>T: 715-723-1181<br>E: jstafford@wasdinet.org | ***Frances Lawrenz***<br>T: 612-625-2046<br>E: lawrenz@umn.edu |
| **Rocky Mountain Middle School** | ***Doris Kimbrough***<br>T: 303-556-3202<br>E: Doris.Kimbrough@CUDENVER.edu<br><br>***Carole Basile***<br>T: 303-556-3336<br>E: carole.basile@cudenver.edu | ***Barbara Heath***<br>T: 910-784-9523<br>E: bpheath@bizec.rr.com<br><br>***Brian Cobb***<br>T: 970-491-6835<br>E: R.Brian.Cobb@ColoState.EDU<br><br>***Bonnie Walters***<br>T: 303-556-6512<br>E : Bonnie.Walters@cudenver.edu |
| **Science Lead Teacher and Science Coalition Courses** | ***Kelli Martin***<br>T: 302-857-3352<br>E: kmartin@doe.k12.de.us | ***Kelli Martin***<br>T: 302-857-3352<br>E: kmartin@doe.k12.de.us |
| **South Carolina Coaching Initiative** | ***John Holton***<br>T: 803-734-8311<br>E: jholton@sde.state.sc.us | ***Yvonne Michel***<br>T: 843-509-5943<br>E: michely@musc.edu |
| **TaLL** | ***Mary Switzer***<br>T: 609-771-2714<br>E: switzer@tcnj.edu<br><br>***Sharon Sherman***<br>T: 609-771-2964<br>E: shermans@tcnj.edu | ***Richard Grip***<br>T: 877-299-6412<br>E: rsg@statforecast.com |
| **TQE/STEP** | ***Melanie Haas***<br>T: 303-352-4996<br>E: mhaas7@mscd.edu | ***Robin Leake***<br>T: 303-477-4896<br>E: robin@jvaconsulting.com |
| **Willamette Valley** | ***Glen Fielding***<br>T: 503-385-4612<br>E: glen.fielding@wesd.org | ***Joe Hansen***<br>T: 541-745-5254<br>E: hansenjb@wou.edu |