

**This report describes the first year of the funded professional development activities in the Title IIB Math and Science Partnership projects in the Northwest Region and the evaluation models. The analysis is structured around the factors of professional development associated with changes in teacher knowledge and practice.**

---

## OVERVIEW

The Title IIB Mathematics and Science Partnership (MSP) program is a federally funded formula grant program to the states that represents a major investment in the professional development of mathematics and science teachers under the No Child Left Behind Act. The program supports partnerships between the mathematics, science, and engineering faculty of higher education institutions and high-need school districts to increase teachers' content knowledge and pedagogical skills. Other partners, including schools of education, business, and nonprofit organizations, may also join the work. In fiscal year (FY) 2003 an appropriation of just under \$100 million supported more than 300 projects. In FY2004 \$150 million was appropriated and awarded to the states in fall 2004 and winter 2005.

Each state was responsible for administering a competitive grants program with the federal funds (box 1). As part of this process, the states established program priorities and a review process to award projects that responded to the most pressing problems in science and mathematics education that could be addressed through professional development. After funding the projects, the states were responsible for monitoring progress and working with the U.S. Department of Education to document their effectiveness. The enabling legislation required states to annually report to the U.S. secretary of education on the programs' impact on teachers' content knowledge and on student learning.

---

This report addresses three research questions

This study is intended to provide policymakers, state agency staff, and university and school personnel interested in potential Title IIB MSP projects with information about how the MSP program has been implemented and evaluated in the Northwest Region in the first cohort of the funded projects. Results of the Northwest Regional Educational Laboratory needs assessment document the growing concern by educators in

## BOX 1

### Details of the Title IIB Mathematics and Science Partnership legislation

The Title IIB MSP legislation provides considerable guidance for how the Title IIB programs in the states should be structured to improve teacher quality in science and mathematics.

#### Five goals for improving teaching

The legislation has five goals for improving the teaching force in K–12 science and mathematics education:

- Improving and upgrading the status and stature of mathematics and science teaching by encouraging higher education institutions to assume more responsibility for improving mathematics and science teacher education through a comprehensive, integrated system of recruiting, training, and advising mathematics and science teachers.
- Focusing on career-long intellectual growth of teachers and upgrading of skills and knowledge.
- Bringing mathematics and science teachers at both the elementary and secondary levels together with scientists, mathematicians, and engineers to increase teachers' subject matter knowledge.
- Developing more rigor in science and mathematics curricula to align them with state and local academic content standards and with the standards expected for postsecondary study in engineering, mathematics, and science.

- Improving and expanding training of mathematics and science teachers, including training in the effective integration of technology into curricula and instruction (Title IIB, Section 2201, Purpose, Definitions).

#### Required partners

The legislation identifies the members of the partnership, including a state education agency; science, engineering, and mathematics faculty at a higher education institutions; and a high-need local education agency. Other partners might include additional science, engineering, and mathematics faculty; teacher preparation faculty; additional local education agencies (public charter or private schools); businesses; and nonprofit or for-profit organizations.

#### Authorized activities

Although the legislation defines one possible model of professional development—the summer workshop or institute—a range of 10 authorized activities are suggested, including:

- Professional development activities that increase mathematics and science content knowledge for teachers.
- Recruitment of mathematics, engineering, and science majors to teaching through a variety of mechanisms.
- Development of rigorous science and mathematics curricula.
- Development of distance learning programs for mathematics and science teachers.

- Design of programs to connect teachers to practicing scientists.
- Development of teachers and programs to encourage women and underrepresented populations in postsecondary study of science, technology, engineering and mathematics careers. (Title IIB, Section 2201, Grants for Mathematics and Science Partnerships).

The definition of a summer workshop or institute includes direct interactions of at least two weeks between teachers, as students, and higher education institution faculty are required. In addition, the workshop or institute must include at least three days of follow-up in the classroom or through distance learning.

#### Evaluation and accountability plan

MSP partnerships are required to develop an evaluation and accountability plan that measures the impact of activities. The legislation clearly articulates that the evaluation needs to focus on the impact of the professional development on student achievement and must include measurable objectives to increase the number of teachers who participate in content-based professional development (Title IIB, Section 2201, Grants for Mathematics and Science Partnerships).

Evaluations might also include measurable objectives to increase participation by students in advanced courses in mathematics and science and to increase percentages of teachers with academic majors or minors in science, technology, engineering, and mathematics or classes taught by such teachers.

the Northwest Region on implementing research-based instructional practices in core subject areas to directly improve student achievement. Respondents recognized that more effort is needed in identifying “research-based” best practices, and a majority said that they are particularly interested in professional development in specific research-based mathematics practices. The No Child Left Behind Act requires science assessments beginning in 2007–08, and 7 of 10 principals indicated that their schools need to put more effort into student proficiency in science. However, responses regarding professional development needs suggest that most schools have not yet given much thought to specific practices or models they might consider.

Three research questions structure this report:

1. What is the nature of the professional development provided by the Title IIB MSP projects in the Northwest Region?
2. What is the nature of the evaluation of the Title IIB MSP projects in the Northwest Region?

3. Under what conditions is the development of experimental or quasi-experimental models of evaluation appropriate and successful?

Descriptive analyses were conducted of the year one Title IIB MSP projects in Alaska, Idaho, Montana, Oregon, and Washington (see appendixes A and B). These analyses are intended to present the projects as they existed at the end of the first year of funding viewed through the lenses of criteria for high-quality professional development and program evaluation. A related study is currently being conducted by the Council of Chief State School Officers in a project funded by the National Science Foundation (box 2).

The analysis is based on the available documents produced by the states and funded projects and an interview structured around core programmatic issues of the professional development and evaluation designs (appendix C). The documents included state requests for proposals, funded proposals, and year one annual reports. These are static documents, collected before the projects unfolded.

#### BOX 2

### **Details of the Improving Evaluation of Professional Development in Mathematics and Science Education project**

A related study is currently being conducted by the Council of Chief State School Officers (CCSSO) in a project funded by the National Science Foundation. The purpose project is to identify professional development programs that meet criteria established by research and to report on the effects of the identified programs on improving teaching and learning in mathematics and science. The CCSSO study convened a review panel to examine

27 nominated programs from 15 states using a program quality rubric to assess the professional development and evaluation designs (Blank, 2006). The preliminary results are available online at [www.ccsso.org/projects/Improving\\_Evaluation\\_of\\_Professional\\_Development/](http://www.ccsso.org/projects/Improving_Evaluation_of_Professional_Development/).

A detailed comparison of the results of the CCSSO review with the results of this study is not appropriate because the two projects are different in both nature and scope. The CCSSO study is not limited to Title IIB MSPs and includes projects funded by the National Science Foundation. Also, the CCSSO study conducted an analysis around

iterative interviews with project staff and evaluators as well as project documents, while this study relies on a static database. Finally, the CCSSO study uses a rating system implemented by an expert panel, while this study is descriptive and uses similar but not identical criteria in its analysis. Although a detailed comparison is not possible, some general statements about the CCSSO findings are provided in this report. Overall, the preliminary findings from CCSSO suggest results similar to those of this analysis—namely, that there is variation among the projects in terms of both professional development and evaluation design.

The nature of the professional development being implemented in each of the projects was examined using an analytic framework based on the National Evaluation of the Eisenhower Professional Development Program (Desimone et al., 2002; Garet, Birman et al., 1999; Garet, Porter et al., 2001; Porter et al., 2000). The framework is organized around six features of high quality professional development that were identified in that evaluation of mathematics and science programs: duration, activity type, collective participation, content focus, active learning, and coherence (see appendix D). There are four dimensions of partnership: supporting preconditions, complexity, interdependence, and communication (Kingsley & O’Neil, 2004; Kingsley & Waschak, 2005).

The analysis of the evaluations of Title IIB MSP projects used a matrix of evaluation criteria developed by SRI International in the Online Evaluation Resource Library based on the program evaluation standards established by the Joint Committee on Standards for Educational Evaluation (JCSEE, 1994). The program evaluation standards focus on the utility, feasibility, propriety, and accuracy of evaluations of educational programs. The Online Evaluation Resource Library web site includes three matrixes that provide a descriptive framework of what should be included in an evaluation that meets the program evaluation standards.

---

#### Title IIB MSP programs vary across the five Northwest Region states

Alaska is the only state among the five that funded only one project in the first year. The Alaska project served 51 teachers in two school districts: Anchorage—a large urban district—and Lake and Peninsula—a large rural district with many remote schools. The partnership for this project included the two districts and the University of Alaska–Anchorage. The Alaska MSP program focused on K–8 mathematics. Some of the requirements identified by the state included a summer institute of at least two weeks with follow-up during the school year, coursework to help teachers achieve

highly qualified status, and the participation of master teachers to serve as instructors.

Idaho funded four MSP projects in year one, all of which focused on mathematics. The total number of teachers served was 163. Two of the projects served schools in northern Idaho and two were based in southern Idaho. Although the Idaho projects all focus on mathematics, the state did not exclude science projects from the Idaho MSP program.

There were six MSP projects funded in Montana. Three projects focused on mathematics, two on science, and one on both. The projects served 165 teachers in the first year. Three projects served teachers in western Montana, one project was based in eastern Montana, and two projects served teachers from both sides of the state.

In year one, Oregon funded four projects. Three projects were focused on mathematics and one on science. The projects served approximately 142 teachers. Two projects were based in northwestern Oregon, one in the central region of the state, and one in the eastern region. The Oregon MSP program emphasized K–12 mathematics, although it did not exclude science-focused projects from eligibility. The state required a two-week institute with follow-up during the year. Unlike other states in the region, Oregon gave preference to projects that used an experimental or quasi-experimental design for the evaluation. Due to these challenging requirements, the Oregon MSP program also recommended that the projects limit the number of teachers served.

Washington funded four projects in year one. Two projects integrated mathematics and science, one project treated both subjects separately, and one project focused only on mathematics. The partnerships served 258 teachers and were distributed across the state—two were based in western Washington, one in central Washington, and one in the eastern region of the state.

The Washington MSP program required a two-week summer institute with follow-up. The state

also tailored the requirements to emphasize Washington priorities for mathematics and science, placing emphasis on pedagogical content knowledge and teacher collaboration.

---

The Title IIB program has multiple implications for analyzing how policy is translated into professional development program implementation

The Title II legislation focuses heavily on policy-making at the federal and state levels around the quality of the teaching workforce (Plecki & Loeb, 2004). Many of the policy issues in Title II focus at the level of the classroom teacher, including recruitment, induction, retention, certification, and compensation. However, the Northwest Region states have established Highly Objective Uniform State Standards for Evaluation, as required by the No Child Left Behind Act, that have resulted in most veteran teachers already being identified as highly qualified.

One policy issue is what states are doing to provide support for teachers' work. As the primary vehicle for professional development in science and mathematics education, the Title IIB MSP program is a key leverage point in state education policy, determining what is sanctioned as effective professional development. Title IIB MSP projects represent one way that states are addressing the issue of highly qualified teachers. For the most part, Title IIB MSP resources are used largely to improve the disciplinary knowledge and teaching skills of in-service teachers, who are typically already licensed to teach. This policy issue is the focus of the first research question articulated below.

A second policy issue is the nature of acceptable evidence of the effectiveness of professional development. This includes a project evaluation that addresses measurable objectives of improvement of teacher mathematics and science content knowledge and pedagogical skills and of increases in student achievement. The U.S. Department of Education's *Strategic Plan 2002–2007* provided criteria to “transform Education into an evidence-based field” (U.S. Department of Education, 2002,

p. 53). The Education Science Reform Act of 2002 emphasizes the role of experimental and quasi-experimental research and evaluation to determine the effectiveness of educational programs. This policy issue is the focus of the second and third research question articulated below.

The purpose of the analysis in this report is to provide a description of the professional development and evaluations conducted in the first year of the Title IIB MSP programs in the Northwest Region. The analysis is not intended to provide information about the effectiveness of the professional development or the evaluations. Such an analysis would require more extensive sources of data and a direct assessment of the projects, which are beyond the scope of this project.

---

### **WHAT IS THE NATURE OF THE PROFESSIONAL DEVELOPMENT PROVIDED BY THE TITLE IIB MSP PROJECTS IN THE NORTHWEST REGION?**

This section summarizes the findings from the descriptive analysis and provides an overview of the professional development conducted in year one. The nature of the professional development being implemented in each of the projects was examined using an analytic framework based on the National Evaluation of the Eisenhower Professional Development Program (Desimone et al., 2002; Garet, Birman et al., 1999; Garet, Porter et al., 2001; Porter et al., 2000). Full descriptions of each project can be found in appendix A. The framework is organized around six features of high quality professional development: duration, activity type, collective participation, content focus, active learning, and coherence. (The analysis framework for professional development is included in appendix D).

In conducting the descriptive analysis of the professional development, the proposals, year one evaluation reports, and interviews were examined to identify evidence related to each of the criteria and subcriteria in the analysis framework. The data were organized into an analysis framework

matrix and then summarized into the project tables that are included in appendix A. The descriptions should not be read as an evaluation of the effectiveness of the projects.

The summary of the qualitative analysis that follows was developed to look at patterns in professional development provided by the Title IIB MSP projects in the Northwest Region. The intent of the summaries is to provide an overview of the information available in the full descriptions.<sup>1</sup>

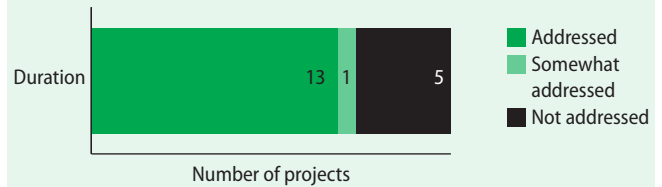
To create this summary, the authors iteratively assigned projects to categories based on the available information related to the dimensions of the analysis framework. The categories that were used to summarize the qualitative analysis included the extent to which the criteria were addressed, or reported. For example, a component of a project was assigned to the “Addressed” category if information related to the dimension was explicitly put forth in both the proposal and the annual report. The “Somewhat addressed” category was assigned if the criterion was treated minimally or the information provided was incomplete. For example, some of the proposals and annual reports included statements that the project was aligned with standards, but no specific information was provided. The category “Not addressed” indicates that there was evidence that the project did not include the aspect of professional development or evaluation, while “Not reported” indicates that the documentation does not provide sufficient information to assign the project to a category.

#### Duration provides opportunities for in-depth study and ongoing support

Duration includes the total number of contact hours and the span of the program. Duration is related to the quality of professional development in at least two ways (Garet, Birman, et al., 1999). First, longer activities provide opportunities for in-depth study of content and pedagogy. In addition, activities that take place over a substantial span of time provide more opportunities to support teachers in trying out new practices. Exemplary professional development programs generally provide

FIGURE 1

#### Most projects provided 80 or more hours of professional development



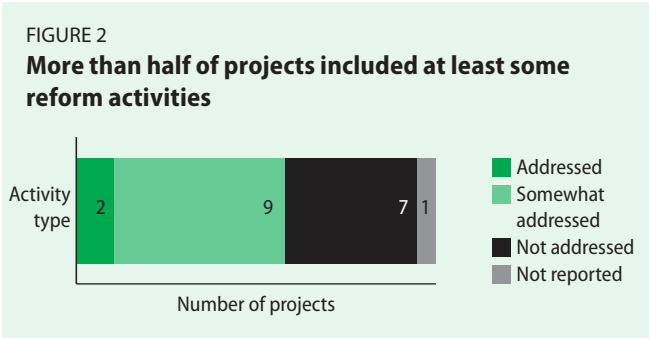
learning experiences that are at least 80 hours long (Garet, Birman, et al. 1999).

In year one 13 projects provided 80 or more hours of professional development and were assigned to the “Addressed” category (figure 1). One project was assigned to the “Somewhat addressed” category because evidence of duration was reported in terms of weeks rather than hours or days. The remaining five projects did not provide sustained professional development.

#### Activity type has two dimensions: traditional and reform

Activity type has two dimensions. Traditional activities include within-district workshops and conferences, courses for college credit, and out-of-district workshops and conferences. Reform activities include teacher study groups; teacher collaboratives, networks, and committees; mentoring; internships; and resource centers. Reform activities often take place during the school day and may be more consistent with teachers’ goals and other professional development activities (Garet, Birman, et al., 1999).

Two projects were assigned to the “Addressed” category because their primary means of providing professional development was reform activities (figure 2). Nine projects were assigned to the “Somewhat addressed” category because they provided a mix of traditional and reform activities. The eight projects in the “Not addressed” category conducted only traditional activities. One project was included in the “Not reported” category because there was not enough information to characterize the activities.



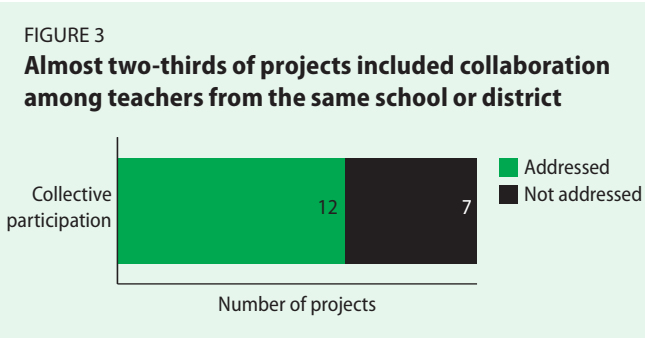
**Collective participation includes activities for teachers from the same school or district**

Collective participation includes activities designed for groups of teachers from the same school or district as opposed to individual teachers. Activities that feature collective participation provide opportunities for teachers to support each other as they attempt to incorporate new knowledge and practices (Garet, Birman, et al., 1999). Collective participation may also make it more likely that structural or organizational support will be put in place.

Twelve projects supported collaboration among teachers from the same school or district and were thus included under the “Addressed” category (figure 3). The remaining seven projects were not designed to explicitly support collective participation, although many of them did provide opportunities for participants to collaborate with each other. While such opportunities may be valuable, they do not address the same issues that are reflected in the criterion for collective participation—fostering schoolwide or districtwide support for teacher growth and improved instruction.

**Content focus addresses the substance of professional development**

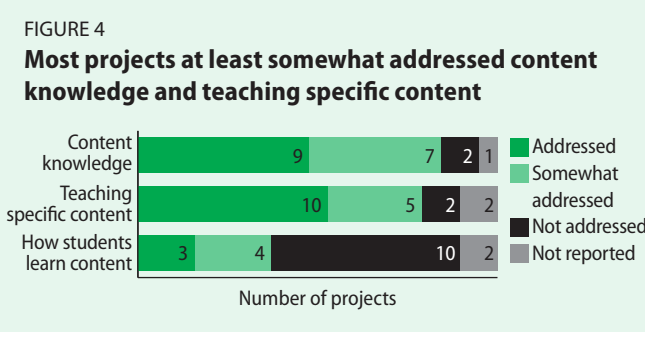
Content focus is the degree of emphasis on deepening teachers’ content knowledge in mathematics and science. There is evidence that professional development focused on content is related to increased student achievement (Cohen & Hill, 1998; Kennedy, 1998). Content focus has three



dimensions: content knowledge, methods of teaching specific content, and emphasis on how students learn specific content (Garet, Birman, et al., 1999; Garet, Porter, et al., 2001).

Projects were assigned to categories for each of the dimensions—some projects addressed multiple aspects of content focus while others addressed only one. Projects that appeared to have a less focused approach—indicated by coverage of multiple topic areas or lack of detail about the content—were assigned to the “Somewhat addressed” category. Two projects in the “Not addressed” category did not identify a focus for the professional development, while one project did not provide any information about the content of the professional development. More projects fell in the “Addressed” category for emphasis on content knowledge and teaching specific content than for how students learn content (figure 4).

In interviews project staff identified a tension between having enough time to address both content and how to incorporate such content into instructional units. However, many project staff also indicated that two weeks was the limit for engaging teachers.



### Active learning describes participants' learning experiences

Active learning includes opportunities for teachers to use new knowledge and practices with support and feedback. It is divided into four dimensions to provide a more specific description of participants' learning experiences: observing and being observed, planning for classroom implementation, analyzing student work, and presenting and leading discussions and writing reports or plans (Garet, Birman, et al., 1999).

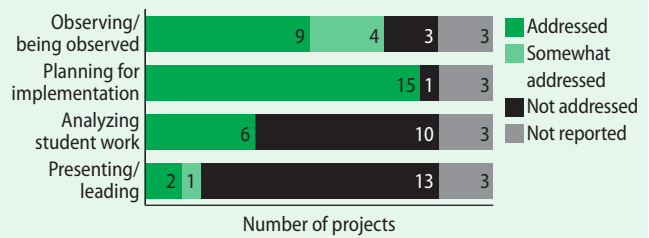
All projects provided some form of active learning except for three that did not provide evidence related to the format of the learning opportunities and were assigned to the category "Not reported." Projects that provided evidence of some types of active learning but not others were placed in the "Not addressed" category for the dimensions not included in their activities. Projects that provided opportunities for participants to observe modeled instruction but not to be observed and receive feedback were categorized as "Somewhat addressed" for the observing and being observed dimension (figure 5); one project provided opportunities for a small subset of participants to present and was categorized as "Somewhat addressed" for the presenting and leading dimension.

The most common form of active learning in the Northwest Region MSP projects was planning for implementation, which was included in 15 projects. Nine projects provided opportunities for participants to observe instruction, to be observed, or both. Significantly fewer projects conducted activities in which participants had opportunities to make presentations or to analyze student work.

In interviews disciplinary faculty indicated that integrating professional development into the work of classroom teachers was a challenge. Lack of sufficient funds to support ongoing interactions with faculty in institutions of higher education and the difficulty integrating service work with schools into tenure requirements were two problems noted.

FIGURE 5

### Most projects provided active learning in the form of planning for implementation



### Coherence indicates how connected professional development is to other teacher learning and development activities

Coherence is the degree to which the professional development is part of a wider set of activities for teacher learning and development. It includes connection to other professional development activities and to teachers' professional goals, alignment with state and district standards and assessments, and support for sustained communication among teachers (Garet, Birman, et al., 1999). Coherence was related to improvements in knowledge and practice reported by teachers (Desimone et al., 2002; Garet, Porter, et al., 2001).

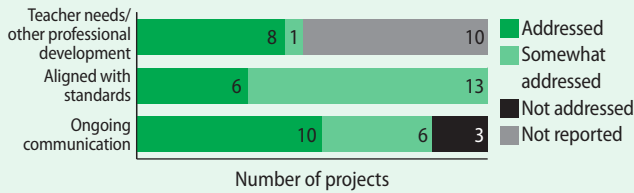
Eight projects provided evidence that the activities were connected to teacher needs or to other professional development and were assigned to the "Addressed" category (figure 6). One project reported providing information that could help create coherence in this dimension and was placed in the "Somewhat addressed" category. The other 10 projects did not provide any information on this dimension.

All of the projects provided some evidence that the activities were aligned with standards, but six projects provided more detailed information and were assigned to the "Addressed" category. Ongoing communication was more common than connecting the project to teachers' needs or other professional development efforts. Projects were assigned the category "Somewhat addressed" if issues of ongoing communication were identified, but it was not clear how the issues would be addressed.



FIGURE 6

**Ongoing communication was the most common dimension of coherence addressed by projects**



**There are four criteria for describing partnership**

The analysis of the professional development also includes criteria that examine the nature of the partnerships that are demonstrated in project documents and through interviews. It has been more problematic to structure the criteria for what constitutes an effective partnership. The work of Gordon Kingsley and his group at the Georgia Institute of Technology has informed the development of the characteristics to describe “partnership” in the Title IIB MSP projects. A partnership, for the purpose of this analysis, is a group of entities (organizations such as schools, colleges or universities, and for-profit or non-profit companies) that work together to accomplish a set of mutual goals. The four dimensions for describing partnership include supporting preconditions, complexity, interdependence, and communication (Kingsley & O’Neil, 2004; Kingsley & Waschak, 2005).

For the summary of partnership descriptions, the projects were divided into only two categories. The “Included” category indicates that the documentation included information related to that dimension. The “Not included” category indicates that the projects did not provide information. This binary system was used because the available evidence related to partnership categories was very limited. In addition, many of the dimensions related to partnership are not indicators of quality but are merely descriptive.

**Preconditions for partnership include existing relationships and mutual goals**

Partnership preconditions include existing relationships between organizations prior to the

development of the partnership. An additional precondition is the extent to which the needs of the partners are congruent and the partnership enables them to pursue mutual goals. Just over half the projects included information about prior collaboration (figure 7). Only three projects included goals that were mutually beneficial to all partners—most of the activities were designed to produce outcomes only for teachers and schools.

**Complexity characterizes the structure of the partnership**

Complexity encompasses several dimensions to describe the structure of a partnership. A partnership with a vertical structure is hierarchical, and a partnership with a horizontal structure includes peer organizations on the same level. Partnerships with sector complexity include organizations with different areas of work. Spatial complexity refers to size of the geographic area that the partnership serves.

The MSP projects were characterized as having either a vertical or horizontal structure, with more projects falling into the vertical category than the horizontal category (figure 8). These categories are merely descriptive—one type of structure is not

FIGURE 7

**Most activities were designed to produce outcomes only for teachers and schools**

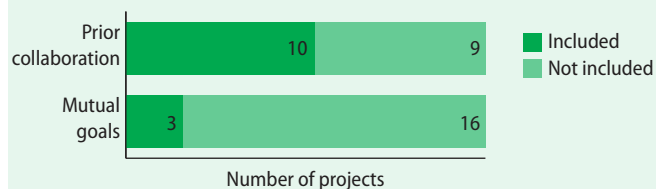
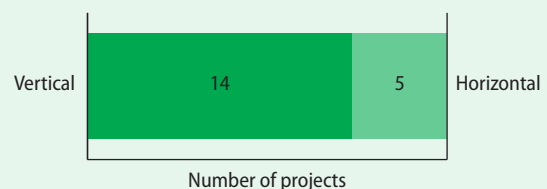


FIGURE 8

**More projects had a vertical structure than had a horizontal structure**



considered more desirable than the other. Only one project involved partners from sectors other than education (figure 9). On the other hand, 10 projects were spatially complex and served a large geographic region.

### Interdependence describes how the partners organize their work

Interdependence characterizes the extent to which partners depend on each other for resources or materials and how they accomplish their work. Partners with reciprocal interdependence share their work back and forth. With sequential interdependence, the work of one partner will feed into the work of another. Pooled interdependence characterizes partnerships in which the members work independently for the most part.

None of the projects indicated that the partners had a sequential approach to interdependence (figure 10). This is likely the result of the available documentation, which provided little detail about how the projects intended to go about partner collaboration. More projects fall into the pooled category, in which the partners have different responsibilities, than into the reciprocal category, in

which the partners share responsibilities. Some of the projects used a mixture of pooled and reciprocal collaboration.

### Communication describes the frequency of contact among partners

The final partnership dimension is communication. The characteristics of communication in the projects were more difficult to identify than those in the other dimensions, and the inferences made are more tentative. Communication refers only to the frequency of communication among the partners rather than the quality or direction of communication. Just over half of the projects provided information about frequency of communication in the documentation (figure 11).

### Only a brief comparison with the Council of Chief State School Officers project is possible

At the time this report was written, CCSSO has not released a full-scale report of its study, *Improving Evaluation of Professional Development in Mathematics and Science Education*. As a result, only a brief comparison between the two projects is possible. Some of the general findings from the CCSSO study include a consistent focus on content knowledge (12 of 15 projects) and alignment with standards (13 of 15 projects). More of the projects in the Northwest Regional Educational Laboratory study provided active learning opportunities via planning for implementation (15 of 19 projects) and analyzing student work (6 of 19 projects) than the CCSSO projects did—7 of 15 projects for planning and 3 of 15 projects for student work.

FIGURE 9

#### More than half the projects served a large geographic region

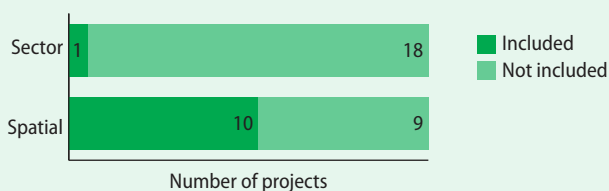


FIGURE 10

#### A pooled approach to collaboration was more common than a reciprocal approach

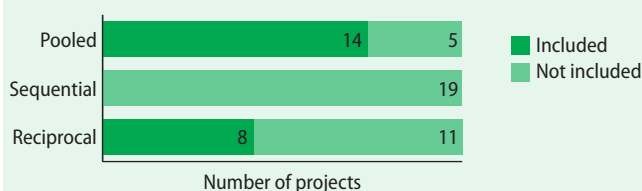
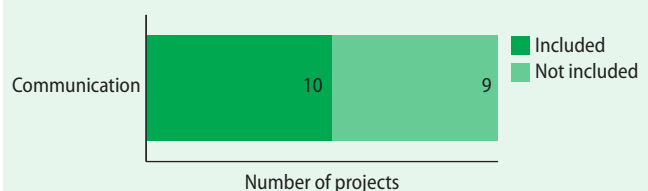


FIGURE 11

#### Just over half of projects provided information about frequency of communication in the documentation



## WHAT IS THE NATURE OF THE EVALUATIONS OF THE TITLE IIB MSP PROJECTS IN THE NORTHWEST REGION?

The second research question examines the nature of the evaluations of the Title IIB MSP projects and the extent to which the project addressed standards for the design and implementation of program evaluations. The authors constructed descriptions of the nature of the evaluation in each project using evidence from the project documents and interviews conducted at the end of year one. The data were organized into a project evaluation analysis framework matrix and then summarized into project tables (see appendix B).

The summary of the qualitative analysis that follows was developed to look at patterns in the evaluations provided by the Title IIB MSP projects in the Northwest Region. The intent of the summaries is to provide an overview of the information available in the full descriptions. The descriptions of the project evaluation should not be read as an evaluation of the effectiveness of the projects. The analysis framework for project evaluation can be found in appendix D.<sup>2</sup>

To create the summaries of evaluation design and implementation, the authors iteratively assigned project evaluation features to categories based on the available information related to the dimensions of the evaluation analysis framework. A feature of a project was assigned to the “Addressed” category if evidence related to the dimension was explicitly put forth in the proposal, the set of project interviews, and the annual report. The “Somewhat addressed” category was assigned if the criterion was treated minimally or the information provided was incomplete. For example, in many of the proposals and annual reports, the evaluator was identified, but no information was provided about that person’s experience to establish credibility. The category “Not addressed” indicates that there was evidence that the project did not include the aspect of professional development or evaluation, while “Not reported” indicates that the documentation does not provide sufficient

information to assign the project to a category. More detail about how categories were assigned is provided in the summary sections for professional development and evaluation in the main report.

### Project context provides information about factors that might influence implementation

A key element of an evaluation should be sufficient information about the project so that the evaluation audience can make inferences about factors that might contribute to project implementation and impact. This criterion includes a description of the project goals, objectives, and activities and an identification of appropriate stakeholders. Evaluations should also include some information about contextual factors that might have influenced the project.

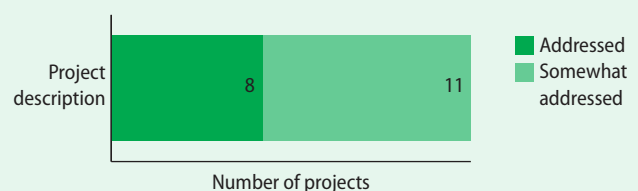
All the projects included some information about the project and identified the appropriate stakeholders in their proposals or annual reports. However, only eight projects included information about the project context that might have influenced project implementation (figure 12).

### Evaluation purpose should be clearly identified

A second criterion of evaluation quality addresses the extent to which the evaluations of the projects identified ways in which information was used. Formative evaluation information should be reported by the extent to which it was used by project staff to improve the professional development. Summative evaluation information communicates the extent to which project goals and objectives have been met.

FIGURE 12

**Less than half of projects included information about the project context that might have influenced project implementation**



A majority of the projects included at least some identification about the purpose of the project, and formative purposes were most frequently cited (figure 13). Only four project evaluations included details about how the projects used the information from the evaluation to improve project function. Almost one-third of the projects did not indicate the purposes the evaluation served.

**Evaluation questions should be aligned with project goals**

A key feature of a quality evaluation is the identification of evaluation questions and their alignment with project goals and objectives. In addition, evaluation questions should indicate how evaluation information might be used by appropriate stakeholders.

Only two projects had evaluation questions that were clearly articulated, well aligned with project goals and objectives, and explicitly tied to data sources (figure 14). Another five projects stated evaluation questions, but the questions were either not aligned with project goals or not connected to sources of information to answer them. In the remaining projects evaluation questions were either not stated simply repeated the project goals or objectives.

**The identity and credibility of the evaluator should be clear**

The evaluator of a project should be clearly identified and the relationship between the evaluator and project should be articulated. In addition, the professional qualifications and experiences of the evaluator should be briefly described to establish one element of trust in the findings.

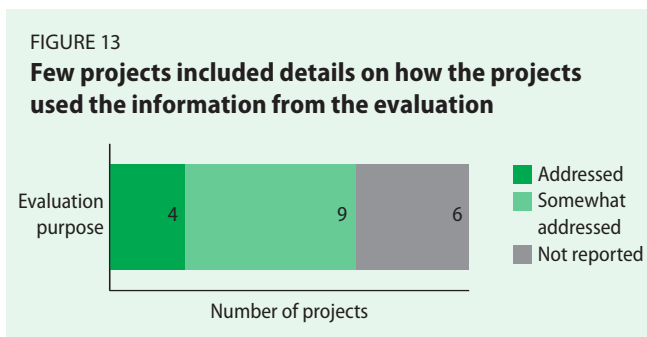
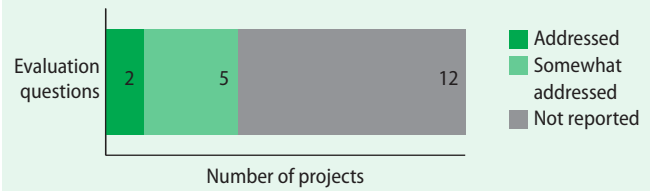


FIGURE 14  
**Few projects had evaluation questions that were clearly articulated, well aligned with project goals and objectives, and explicitly tied to data sources**



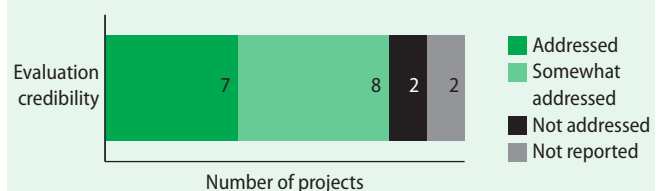
Seven projects identified a specific person for the evaluation and included information about the evaluator’s qualifications (figure 15). In eight projects the evaluator was identified, but no additional information about the evaluator was included. In two projects the evaluator identified in the proposal was not used for the evaluation, and in two others the evaluator was not identified at all.

**Evaluation plans should describe stakeholder involvement**

An evaluation should describe how the positions and perspectives of the stakeholders will be taken into account throughout the evaluation. A description of stakeholders’ involvement in the evaluation illustrates how they influence the evaluator’s understanding of project goals and objectives, how evaluation questions are shaped, and how results are reviewed.

Only seven project evaluations mentioned stakeholder involvement with the evaluation (figure 16). Involvement was defined as having contributed information to the evaluation, and formative use of evaluation information was largely limited to

FIGURE 15  
**More than a third of projects identified a specific person for the evaluation and included that person’s qualifications**



project staff responsible for the development of professional development activities. In the remaining projects no indication of stakeholder involvement was provided.

**Methodological approaches need to be clearly described**

The evaluation plan should describe the proposed methodological approaches. The description should provide information about how the evaluator will identify and use data that help answer the evaluation questions within the constraints of time and cost. The data gathered need to be aligned with the goals that the project is intended to achieve.

Four projects included sufficient information about the design of the evaluation to identify the methodological approach (figure 17). They attempted some form of experimental or comparison group design, with enough information to explain how the design was implemented. In seven projects a methodological approach was initially identified and included both quantitative and qualitative evaluation designs, but there was insufficient evidence in the annual report that the approach had actually been implemented. Two

projects indicated in their proposals that they would use a comparison group approach but did not implement such a design. The evaluations that used a qualitative evaluation approach included limited information about the design elements. Almost half of the projects did not indicate any methodological approach.

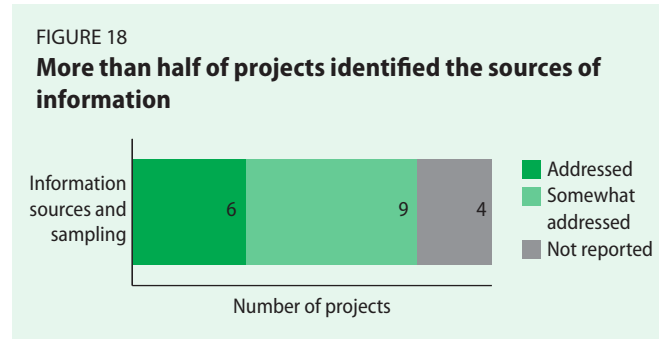
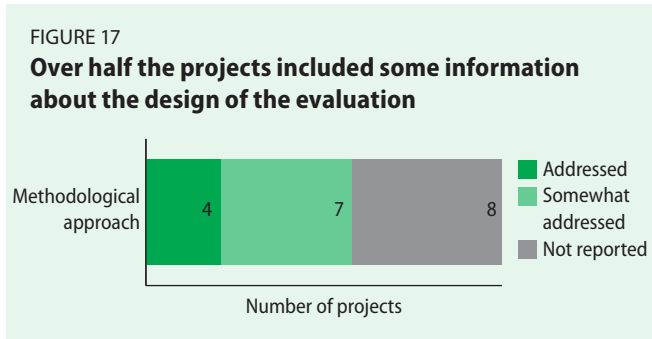
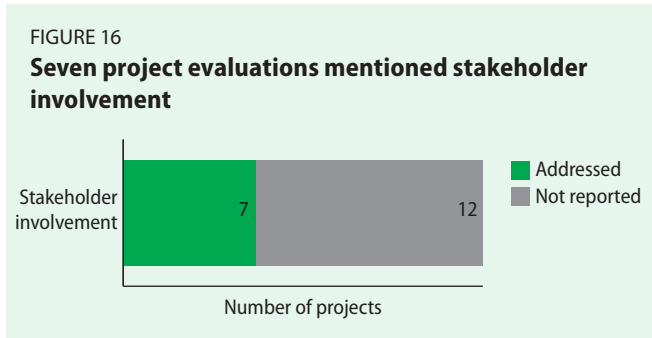
**Plans should report information sources and sampling**

The sources of information that will be used in the evaluation should be described in enough detail to assure that the information will be sufficient to meet the evaluation’s purposes. Information about how the groups were selected to provide information should be identified and briefly described. If stakeholders are sampled, sampling strategies need to be clearly described.

Almost one-third of the project evaluations included some characterization of the sources of information for their evaluations as well as how those sources were chosen and how samples were determined (figure 18). Nine projects identified the sources of information. But four did not include any information about data sources or about who supplied comments about professional development activities, meaning the reader does not know how many participants commented in support of the project.

**Information about data collection instruments is essential**

The evaluation should describe the nature of the data collection instruments and how they will be used to gather needed information. Information should also be included that demonstrates how



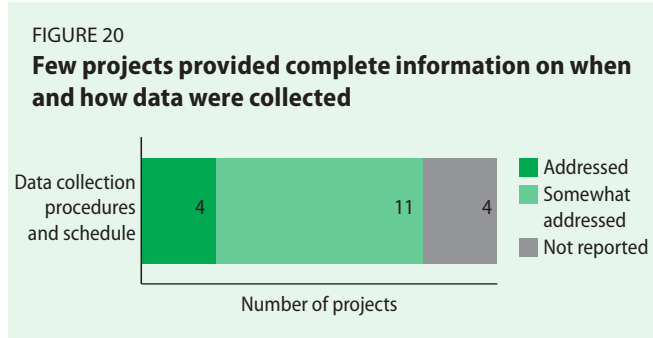
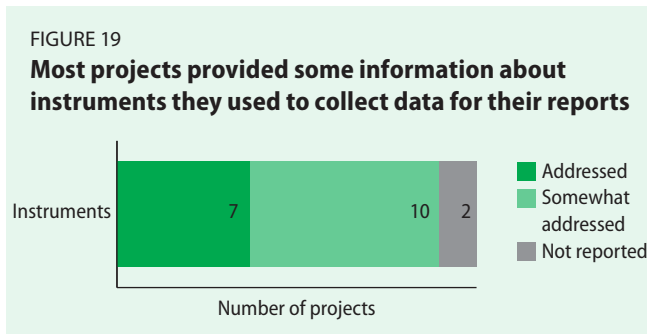
instruments are used to examine evidence of data reliability and validity.

More than one-third of the projects specifically identified the instruments used to collect data and included those instruments as appendixes in their annual reports, but they did not provide information about quality control characteristics of the use of the instruments, such as evidence of reliability and validity (figure 19). An additional 10 projects indicated that they used instruments, frequently project designed, but did not include the instruments in their reports. Two projects did not specify instruments in enough detail to inform the reader of what data were collected.

**Specific data collection procedures should be identified**

The evaluation report should describe how and when data were collected from sources. This information should include how the sources will answer evaluation questions through triangulation and multiple perspectives. The description of the data collection and its intent will provide a context for the eventual judging and interpreting of evaluation findings and recommendations.

Only four projects provided specific information about when and how data were collected, but it was unclear who was responsible for collecting the data (figure 20). A majority of projects indicated that data collection occurred but provided no specific information about when and how. Almost one-quarter of the projects did not address when and how data were collected.



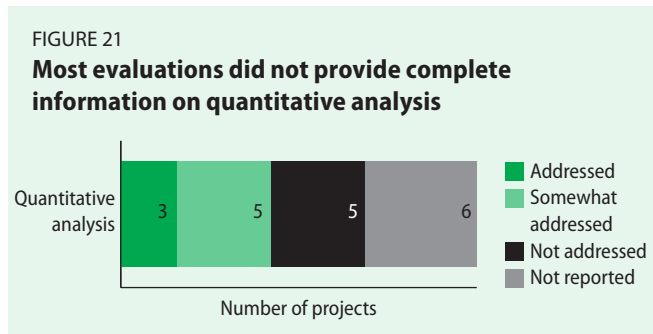
**Descriptions of quantitative analysis processes should be detailed**

Evaluations need to show the nature and appropriateness of the quantitative analysis procedures and their relationships to the evaluation questions and data sources. Information about how the evaluation addresses the practical significance (e.g., effect sizes) and replicability, as well as statistical significance, should be included.

Only three projects included information about which quantitative analysis procedures were used to conduct statistical analysis of project data (figure 21). One-quarter indicated using descriptive statistics to report survey results, and another quarter reported statistically significant results but provided no information about how their analyses were performed. Reporting significance without providing transparent data and analysis processes is a major flaw in these evaluations. The remaining six projects provided no indication of quantitative analysis.

**Qualitative analysis processes should be articulated**

Evaluations need to show the nature and appropriateness of the qualitative analysis procedures and



**FIGURE 22**  
**Only one project provided any information about how qualitative data were analyzed**



their relationships to the evaluation questions and data sources. The extent to which the findings are supported by qualitative data gathered from more than one source should be addressed.

Only one project provided at least some information about how qualitative data were analyzed and addressed triangulation methods (figure 22). Most projects indicated that qualitative data were analyzed but failed to provide sufficient information as to how. And two projects did not report any qualitative data.

**Results from the evaluation analyses here cannot be compared with those from the Council of Chief State School Officers**

The results of the evaluation analyses from this Fast Response Task and the CCSSO study are not comparable because two different sets of criteria were used. The CCSSO study rated the evaluations according to Guskey’s model for evaluating professional development (Guskey 2000), while the Northwest Regional Educational Laboratory study used criteria based on the Joint Committee on Standards for Educational Evaluation (JCSEE 1994).

**UNDER WHAT CONDITIONS IS THE DEVELOPMENT OF EXPERIMENTAL OR QUASI-EXPERIMENTAL MODELS OF EVALUATION APPROPRIATE AND SUCCESSFUL?**

The third research question of this Fast Response Task asks under which conditions experimental or quasi-experimental evaluation designs are

successful or appropriate or successful for the Title IIB MSP projects in the Northwest Region. No project conditions supported successful implementation of an experimental or quasi-experimental evaluation design in the Title IIB MSP projects. But the extent to which such rigorous evaluation models are appropriate for Title IIB MSP projects is less straightforward.

Only 7 of 19 funded projects included some evaluation design that included a control or comparison group. Two projects used randomized control experiments with random assignment of schools to a treatment and control group. Characterizing the other project evaluation designs in a quasi-experimental category is problematic because the evaluators did not provide sufficient indication of how they compared treatment and comparison groups. These projects are thus referred to here as comparison group evaluation designs. Of the five projects with comparison group evaluation designs, two included the comparison group design in the proposal, but there was no evidence of the results in the year one report. Two other projects used a comparison group design but lacked enough detail regarding the identification of the comparison group to determine the utility of the information provided. A fifth project allowed for self-selection of teachers into a treatment or comparison group and provided limited information about the comparison group.

The first issue that arises in the analysis is why only 7 of 19 funded projects even attempted a comparison group, quasi-experimental design, or experimental evaluation design. In interviews project staff and evaluators reported several issues that arose as they initially examined the potential to design and implement rigorous evaluations. The primary reason that projects did not seek to implement an experimental, quasi-experimental, or comparison group design was the difficulty or reported impossibility of recruiting a control or comparison group within the context of the professional development partnership. Project staff indicated that schools and districts refused to consider such assignment. Even when a project

sought to recruit a comparison group, they had difficulty matching the treatment and comparison groups at the grade level when their projects addressed a broad span of teachers. Another reason was the lack of availability of adequate instrumentation. Project staff indicated that even if they had negotiated an agreement for a randomized design, they did not have adequate measures of changes in teacher content knowledge and instructional practices or of student achievement in project-focused content. They indicated that they did not consider the state assessments to be aligned with the content of the professional development.

Adequate evaluation resources were identified as problematic by several projects. In many cases the evaluation budget was reported to be between \$3,000 and \$10,000. Conducting a rigorous evaluation for such a sum was cited as being impossible. Finally, evaluation capacity was indicated as a major reason why the models of evaluation were not rigorous. Few of the evaluators indicated that they had experience with or capacity to carry out a rigorous evaluation.

Two of the projects attempted to implement an experimental design. In both the initial model of evaluation posited random assignment to treatment and control groups at the teacher level. However, both projects had to change that random assignment to a school level as teacher collaboration in either a formal lesson study approach or across grade levels through electronic support were an integral part of the professional development. Changing the unit of analysis from teacher to school resulted in a severe reduction in sample size. Project staff indicated that they were aware of the lack of power in their evaluation design and that this lack of power made determining an effect of the project on teacher content knowledge, instructional practices, or on student achievement very problematic.

Both of the projects that did attempt more rigorous evaluations came from Oregon. The request for proposals in that state indicated that preference would be given to grant applicants who included

randomized controlled trials or similar quasi-experimental methods in the project evaluations. In addition, the bidders sessions provided by the state education staff who were responsible for explaining the grant program to potential applicants emphasized the need to conduct research within the Title IIB MSP program. This is not to say that simply exhorting projects to implement rigorous evaluations will result in projects engaging in such designs. However, the state education agency staff member indicated that including the increased preference for experimental or quasi-experimental evaluations was a major factor in the resulting attempts.

In addition, the U.S. Department of Education has provided continued technical assistance to the state education agency staff and to projects on a regional basis. This technical assistance included information about assessments for both teachers and students that several project staff and evaluators indicated provided them with instruments that they might use to measure change in teacher and student content knowledge. State education agency staff have shared this information with project staff in periodic meetings throughout the past two years. They indicated that projects intended to increase the rigor of their evaluations in the second year of the projects.

---

## FINDINGS SUGGEST FUTURE DIRECTIONS FOR RESEARCH

Two difficulties were encountered in the design of Title IIB MSP professional development programs. One is providing content to teachers and addressing instructional needs. In interviews project staff identified tension between having enough time to address content and the ways in which such content should be incorporated into instructional units, especially within the two week window usually allotted for training workshops. Another issue is the difficulty that project partners, especially disciplinary faculty, have integrating professional development into the work of the teacher in the classroom. Lack of sufficient funds



to support ongoing interactions between faculty in institutions of higher education and the difficulty integrating service work with schools into tenure requirements were two problems noted.

The evaluation designs of Title IIB MSP programs are not rigorous enough to provide gold standard evidence of their effectiveness. Even evaluations that attempted an experimental or comparison group design suffered from common flaws. Evaluations of many projects do not address key measurable outcomes or impacts of the projects. Few projects used well developed instruments to measure changes in teacher content knowledge. In addition, the state assessment systems on which many of the projects relied were inadequate to directly measure the impact of projects on student achievement.

Ongoing technical assistance is needed to increase the evaluation skills of the state education agency staff responsible for the Title IIB MSP programs and the staff and evaluators of the individual projects. The U.S. Department of Education regional forums can be a start in informing stakeholders about the method and instruments to improve evaluations, but access to these regional forums is limited by project budget constraints.

But this analysis is based on the first year of implementing the Title IIB MSP programs, when evaluation designs may not be fully mature. The data analyzed were static, showing what the projects intended to do in the proposals and what they could demonstrate in the first year annual report. The single interview conducted with the staff and evaluator in each project captured information just after the second year of the summer institutes. State education agency staff indicated on several occasions that the projects had learned from their first year and that the projects' professional development changed from year one to year two.

In addition, the analysis was conducted by the two authors with assistance from a policy expert at the Northwest Regional Educational Laboratory. There has been no opportunity to include additional reviews to examine the data sources to refine the analyses.

The Northwest Regional Educational Laboratory's Center for Classroom Teaching and Learning has proposed extending this research to a second year of analysis of the Title IIB MSP projects. Year two annual reports will be collected and an expanded evaluation rubric will be constructed that examines additional elements of the Joint Committee on Standards for Educational Evaluation that addresses the criteria for the use of instruments and evaluation reports. In addition, this second year of the research would send the current analyses by state back to the state education agency staff and through them to the original project stakeholders for a member check. This is not seeking validation of the analysis. Rather, this member check would be an additional data source that would be analyzed to determine how the Title IIB MSP is being manifested in the Northwest Region.

---

## NOTES

1. Presenting this information in this compressed format means that there is a risk that the findings will be misrepresented. To understand what the summaries mean, it is necessary to interpret them in the context of the full project descriptions.
2. Presenting this information here in a compressed format creates the risk that the findings will be misrepresented. To understand what the summaries mean, they must be interpreted in the context of the full descriptions.