



What's Happening

March 2016

# Teacher evaluation and professional learning: Lessons from early implementation in a large urban district

**Karen Shakman**

Education Development Center

**Jacqueline Zweig**

Education Development Center

**Candice Bocala**

WestEd

**Natalie Lacireno-Paquet**

WestEd

**Jessica Bailey**

Education Development Center

In collaboration with the Northeast Educator Effectiveness Research Alliance



**ies** NATIONAL CENTER FOR  
EDUCATION EVALUATION  
AND REGIONAL ASSISTANCE

Institute of Education Sciences  
U.S. Department of Education

**REL**  
NORTHEAST  
& ISLANDS

Regional Educational Laboratory  
At Education Development Center, Inc.

**U.S. Department of Education**

John B. King, Jr., *Acting Secretary*

**Institute of Education Sciences**

Ruth Neild, *Deputy Director for Policy and Research*  
*Delegated Duties of the Director*

**National Center for Education Evaluation and Regional Assistance**

Joy Lesnick, *Acting Commissioner*  
Amy Johnson, *Action Editor*  
Elizabeth Eisner, *Project Officer*

REL 2016–115

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

March 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0009 by Regional Educational Laboratory Northeast & Islands administered by Education Development Center, Inc. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Shakman, K., Zweig, J., Bocala, C., Lacireno-Paquet, N., & Bailey, J. (2016). *Teacher evaluation and professional learning: Lessons from early implementation in a large urban district* (REL 2016–115). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

## Summary

Policymakers and researchers increasingly recommend aligning educator evaluations and professional development to improve instruction and student learning. However, few empirical studies have examined the relationship between new educator evaluation systems and the professional development in which teachers engage following their evaluation. Thus, although the new evaluation systems provide a wealth of data about teacher performance, little is known about how these data are used to address teachers' professional needs. This study, conducted by the Regional Educational Laboratory (REL) Northeast & Islands in collaboration with the Northeast Educator Effectiveness Research Alliance, looked closely at one large urban district's educator evaluation system from May 2013 to May 2014. The study examined the written feedback evaluators provided to teachers who were rated less than proficient in one or more standards of effective teaching practice: curriculum, planning, and assessment (standard 1); teaching all students (standard 2); family and community engagement (standard 3); and professional culture (standard 4).

This written feedback, which the district refers to as prescriptions, suggests how teachers may improve their practice by participating in professional activities, including professional development activities and professional practice activities. Professional development activities involve interaction between the teachers and the evaluator, colleagues, mentors, coaches, or instructors. Workshops, coaching, and formal meetings with colleagues are examples of the professional development activities prescribed. Professional practice activities involve independent work in or outside the classroom through which a teacher may improve his or her practice. Trying different instructional strategies or submitting lesson plans are examples of professional practice activities prescribed.

The study examined which standards teachers received prescriptions for, what type of prescriptions they received, the type of professional activities they reported participating in related to each standard, the extent to which these activities aligned to those prescribed by their evaluators, and whether teachers' ratings improved in a subsequent evaluation. The data for this study consisted of teacher characteristics and evaluation ratings, prescriptions, a district-administered teacher survey in which teachers reported on their professional activities related to each standard, and a small number of teacher and principal interviews.

The following are the key findings from the study:

- Teachers received prescriptions across all four standards, usually for one or two professional activities per prescription, and they received more prescriptions with professional practice activities than with professional development activities.
- Teachers reported participating in more professional activities, including both professional development and professional practice activities, for instruction-based standards (standards 1 and 2) than for non-instruction-based standards (standards 3 and 4).
- For all standards, less than 40 percent of teachers who responded to the survey participated in all the activities their evaluators prescribed. However, at least 75 percent of teachers who received a prescription for standard 1 (curriculum, planning, and assessment) or standard 2 (teaching all students) and responded to the survey reported participating in at least one professional activity that related to those standards. For standards 3 and 4, fewer teachers engaged in the prescribed

activities, but many engaged in other types of professional activities, including professional development or professional practice activities, related to the standard.

- Of the teachers rated less than proficient who had received a prescription for standard 1 and then participated in any professional activities related to that standard, 64 percent received at least a proficient rating on a subsequent evaluation; 34 percent of the teachers with prescriptions for standard 1 who did not participate in related activities also raised their summative rating to proficient. Standard 1 was the only standard for which a statistically significant difference was detected in the subsequent evaluation of teachers who engaged in activities aligned to their prescriptions and those who did not. The percentage of teachers in the study group who received at least a proficient rating on their subsequent evaluation did not vary by whether they participated in the particular type of activity their evaluator prescribed.

While this report does not examine the quality of the feedback teachers received or make causal claims about the impact of professional activities on their practice, lessons from this district's early efforts to align evaluation and support may be valuable to other states and districts embarking on similar work.

## Contents

<b>Summary</b>	<b>i</b>
<b>Why this study?</b>	<b>1</b>
Educator evaluation in the study district	1
Research questions	3
<b>What the study found</b>	<b>6</b>
Teachers received prescriptions across all four standards that tended to include one or two professional activities per prescription	6
Teachers received more prescriptions with professional practice activities than with professional development activities for all standards	7
More teachers reported participating in professional activities, including both professional development and professional practice activities, for standards 1 and 2 than for standards 3 and 4	9
For all standards, less than 40 percent of teachers participated in all the activities their evaluators prescribed	11
A higher percentage of teachers who participated in activities related to curriculum, planning, and assessment received at least a proficient rating in this standard on their subsequent summative evaluation than did teachers who did not participate	13
<b>Implications of the study findings</b>	<b>14</b>
<b>Limitations of the study</b>	<b>15</b>
<b>Appendix A. Extended literature review</b>	<b>A-1</b>
<b>Appendix B. Sample prescription</b>	<b>B-1</b>
<b>Appendix C. Data and methodology</b>	<b>C-1</b>
<b>Appendix D. Coding dictionary</b>	<b>D-1</b>
<b>Appendix E. Supplementary tables</b>	<b>E-1</b>
<b>Appendix F. Survey items</b>	<b>F-1</b>
<b>Appendix G. Interview protocols</b>	<b>G-1</b>
<b>References</b>	<b>Ref-1</b>
<b>Boxes</b>	
1 Key terms	2
2 The Northeast Educator Effectiveness Research Alliance	3
3 Data and methods	5

## Figures

1	Percentage of teachers in random sample who were prescribed one, two, or three or more types of activities, by standard, May 2013–February 2014	7
2	More survey respondents reported participating in any professional activities related to the instruction-based standards (standards 1 and 2) than professional activities related to the other two standards, May 2013–May 2014	9
3	For all standards, less than 40 percent of survey respondents reported participating in all activities their evaluators prescribed, May 2013–May 2014	11
4	For standard 2 a higher percentage of teachers reported participating in at least one prescribed professional development activity than at least one prescribed professional practice activity; the reverse was true for standards 3 and 4, May 2013–May 2014	12
5	A significantly higher percentage of teachers who participated in professional activities related to standard 1 received at least a proficient rating on their 2013/14 summative evaluation than did teachers with a prescription for standard 1 who did not participate in any professional activities related to the standard	14

## Tables

1	Activities associated with professional development and professional practice activities	4
2	Percentage of teachers in random sample prescribed each type of professional activity for each standard, May 2013–February 2014	8
3	Percentage of survey respondents reporting participation in each type of professional activity for each standard, May 2013–May 2014	10
C1	Characteristics of teachers who received a prescription, May 2013–February 2014	C-2
C2	Sample of teachers and number of prescriptions for each research question, May 2013–February 2014	C-3
C3	Comparison between the population of teachers who received a prescription and the stratified random sample of 148 teachers, May 2013–February 2014	C-5
C4	Teachers in the standard populations and in the random sample who received a prescription, by standard, May 2013–February 2014	C-6
C5	Characteristics of teachers who received a prescription, May 2013–February 2014, who responded and did not respond to the survey	C-7
C6	Teachers' evaluation ratings and number of prescriptions for teachers who responded and did not respond to the survey, May 2013–February 2014	C-8
D1	Coding dictionary	D-1
E1	Percentage of survey respondents who reported participating or not participating in professional activities related to each standard, May 2013–May 2014	E-1
E2	Percentage of survey respondents who reported participating in each type of professional activity, by standard, May 2013–May 2014	E-1
E3	Percentage of survey respondents who reported participating in any professional activities in the standard in which they received a prescription, May 2013–May 2014	E-2
E4	Percentage of survey respondents who reported participating in the type of professional activities prescribed by their evaluators, by standard, May 2013–May 2014	E-2
E5	Percentage of survey respondents who reported participating in at least one professional activity prescribed by their evaluators, by standard, May 2013–February 2014	E-2
E6	Percentage of survey respondents who received a prescription, May 2013–February 2014, and were rated at least proficient on the 2013/14 summative evaluation, by standard	E-3
E7	Percentage of survey respondents who were rated at least proficient on their 2013/14 summative evaluation by whether they participated in any professional activities related to the standard for which they had a prescription, by standard, May 2013–May 2014	E-3

E8 Percentage of survey respondents who were rated at least proficient on their 2013/14 summative evaluation by whether they participated in the prescribed types of professional activities in the standard for which they had a prescription, by standard, May 2013–May 2014

E-3

## Why this study?

Policymakers and researchers increasingly recommend aligning educator evaluation and professional development systems to improve instruction, student learning, and school capacity (Coggshall, Rasmussen, Colton, Milton, & Jacques, 2012; Curtis & Wiener, 2012; Danielson, 2012; Darling-Hammond & Stanford Center for Opportunity Policy in Education, 2012; Goe, Holdheide, & Miller, 2014). Some studies affirm that teachers who participate in evaluation systems that provide timely, helpful feedback, and in which teachers perceive the evaluators as supportive, are more likely to seek professional development following the evaluation and to view the evaluation process favorably (Delvaux et al., 2013; Milanowski & Heneman, 2001; Taylor & Tyler, 2011; Tuytens & Devos, 2011, 2014).

Few empirical studies have examined the explicit relationship between the data generated from these new evaluation systems and the subsequent professional activities (see box 1 for definitions of key terms) in which teachers who were rated less than proficient engage (Hamilton et al., 2014). Research has yet to uncover how principals and others use the evaluation systems to guide teachers' continuous professional improvement. (A review of the literature is in appendix A.)

This study, conducted by REL Northeast & Islands in collaboration with the Northeast Educator Effectiveness Research Alliance (box 2), examined how a large urban district's newly developed evaluation system addressed the needs of teachers who were deemed less than proficient. The study further examined the extent to which the feedback evaluators provided to these teachers aligned with the types of professional activities in which teachers engaged as well as how both teachers who did and those who did not subsequently engage in professional activities fared on their next evaluation.

The system examined was in the early stages of aligning evaluation and support for teachers deemed to need improvement in one or more areas. It had established limited oversight of the development and monitoring of the prescriptions. This study does not provide information on the quality of the feedback teachers received or make causal claims about the relationship between participating in professional development or other professional activities and changes in teacher practice. Lessons learned from a district in the early stages of implementation may be valuable to other states and districts as they embark on aligning systems of evaluation and support to efficiently and effectively improve teachers' practice.

### Educator evaluation in the study district

The study district introduced a new educator evaluation system in school year 2012/13 that was designed to allow evaluators to suggest professional resources to support teachers' professional growth. This study examined this system in its second year of implementation.

In the system, a teacher is assessed as exemplary, proficient, needs improvement, or unsatisfactory in each of four standards of effective practice:

- Curriculum, planning, and assessment (standard 1)
- Teaching all students (standard 2)
- Family and community engagement (standard 3)
- Professional culture (standard 4)

***This study examines how a large urban district's newly developed evaluation system addressed the needs of teachers who were deemed less than proficient and the extent to which the feedback evaluators provided to these teachers aligned with the types of professional activities in which teachers engaged***

---

## Box 1. Key terms

**Alignment.** The extent to which teachers' professional activities corresponded to those specified in their evaluation prescriptions. Alignment is categorized as one of the following: fully aligned (all types of activities prescribed were done), partially aligned (at least one of the prescribed types of activities was done, but not all), or not aligned (none of the types of activities prescribed were done).

**Formative and summative teacher evaluations.** Educators in the district examined in this study receive both formative and summative evaluation ratings on the four standards. These ratings are exemplary, proficient, needs improvement, and unsatisfactory. Ratings are assigned to each standard in the evaluation, and evaluators provide a rationale to support each rating. Formative evaluations are completed by February 1 and summative evaluations by May 15 of each year.

**Evaluator.** Evaluators of teachers include school administrators, such as a principal. Evaluators play a large role in the educator evaluation system, including conducting formative and summative evaluations, determining overall effectiveness ratings, and assigning prescriptions.

**Prescriptions.** In the evaluation system examined in this study, evaluators provide written directives to a teacher that indicate professional development or practice activities that could help the teacher improve in any of the four standards in which he or she is rated less than proficient. There is no direct consequence to not fulfilling the prescription if the teacher's practice improves. However, a teacher who receives two consecutive less-than-proficient ratings on standards 1 and 2 at the summative evaluation receives an unsatisfactory rating overall, and more formal action is taken that includes a timeframe for possible dismissal. For standards 3 and 4, teachers have more time to improve their ratings.

**Professional activities.** Professional development activities and professional practice activities designed to improve teacher performance.

**Professional development activities.** As defined in this study, professional development involves interaction between the teacher and the evaluator, colleagues, mentors, coaches, or instructors. It includes traditional professional development (such as courses and workshops) and integrated professional development (such as coaching, meeting with an evaluator or colleagues, and observing colleagues). This study categorized the following activities as professional development: a workshop or course (for example, a school or district workshop during the academic year, a summer workshop, a face-to-face course located outside the district, an online course or webinar, a conference), regular meetings with an evaluator, formal coaching or mentoring by a nonevaluator, formal meetings with colleagues (for example, data team meetings, professional learning community), and observation of a colleague.

**Professional practice activities.** As defined in this study, professional practice activities refer to the independent work in which teachers may engage to improve their practice, either by enacting strategies in the classroom or for their own professional growth outside the classroom. These activities are not typically conducted through interaction with other professionals. This study categorized the following activities as professional practice: document submission (for example, lesson plans, data, artifacts), reading resources recommended by the evaluator, specific instructional strategies recommended by the evaluator, and other professional strategies (for example, interactions with school staff, parents, and community members; meeting deadlines; punctuality; attendance; implementing an action plan).

**Standards.** Standards are the broad categories of knowledge, skills, and performance detailed in the state's regulations. The district's rubric of effective practice includes four standards on which teachers are assessed: curriculum, planning, and assessment (standard 1); teaching all students (standard 2); family and community engagement (standard 3); professional culture (standard 4).

---

---

## Box 2. The Northeast Educator Effectiveness Research Alliance

The Northeast Educator Effectiveness Research Alliance is made up of state and district leaders focused on educator effectiveness and the use of new educator evaluation systems to promote improvements in teaching and learning. Its goal is to provide research to support state and district educator evaluation systems and to build their capacity to evaluate their own systems. One of three core topics in the alliance's research agenda is the relationship among evaluation systems, teacher professional development, and teacher practice. Alliance members' interest in this relationship is motivated by their expectation that the new educator evaluation systems will yield improvements in teaching practice and student outcomes. To achieve this goal, members are working to build coherent systems of evaluation and support to improve teaching performance by identifying teachers' needs and providing professional development that increases teacher effectiveness in the classroom. In particular, members want to understand the extent to which needs identified in the evaluations are addressed through professional development and, ultimately, whether participation in professional development that aligns to the identified needs yields improvement in teachers' subsequent evaluations. This study begins to answer these questions by presenting descriptive information from one large urban district in the REL Northeast & Islands Region.

---

Teachers rated as needs improvement or unsatisfactory on any standard are prescribed professional activities to improve their practice. The evaluator (principal or designated staff member) can submit one or more prescriptions at any time during the school year. Each prescription indicates:

- The standard in which the teacher is expected to improve.
- A problem statement that outlines the issue that has been identified.
- Evidence associated with the issue.
- Prescribed activities for the teacher.

An example of a full prescription is shown in appendix B.

The study team identified nine types of professional activities that evaluators prescribed, and placed them in two general categories: professional development and professional practice. These categories are defined in box 1 and the activity types and examples associated with each category are shown in table 1. Data collection and analysis are described in box 3 and appendix C.

The district's use of prescriptions to align teacher evaluations and recommended activities for teachers who have been identified as needing improvement is similar to approaches used by many other states and districts that require evaluators to provide recommendations or develop action plans for or with teachers who need support to receive a rating of proficient.

### Research questions

The study posed the following research questions:

- In what standards did teachers receive prescriptions?
- For each standard, what professional activities did evaluators prescribe?
- What professional activities did teachers report they had participated in?

**Table 1. Activities associated with professional development and professional practice activities**

Type of professional activity	Description/examples
<b>Professional development</b>	
Workshop or course	Taking any workshop, course, or other class in any content area or instructional method, face-to-face or online.
Meeting with evaluator	Meeting directly with the evaluator, with or without a review of materials during the meeting.
Formal coaching or mentoring by a nonevaluator	Working with a colleague named as a coach or mentor.
Formal meeting with a colleague	Working with a colleague who is not a coach or mentor, for example, other teachers, a content specialist, or a classroom assistant.
Observation of a colleague	Observing a colleague teach or demonstrate an instructional or classroom environment strategy.
<b>Professional practice</b>	
Document submission	Submitting evidence or documentation, such as lesson plans, time logs, or materials, to the evaluator.
Reading resources	Reading a book, article, website, or other resource recommended by the evaluator. Specific recommendations for how to implement the content of the reading may or may not be offered.
Instructional strategies	Using a strategy that relates to student learning, classroom instruction, or classroom environment. For example, “think, pair, share” or “have students work in small groups.”
Other professional strategies	Using a strategy that does not directly relate to the work in the classroom, and may relate to other adults in the building or in the community. For example, “write a parent newsletter,” “keep a parent contact log,” or “arrive to school on time.”

**Note:** The categories of professional development and professional practice are drawn from the relevant literature (see appendix A for an explanation of the two categories).

**Source:** Authors’ analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, evaluator prescriptions, a teacher survey, and interviews.

- Did the professional activities in which teachers participated align with the standards for which they received prescriptions and the activities that were prescribed?
- Among teachers who participated and those who did not participate in any professional activities in the standard in which they had a prescription, what percentage improved their summative ratings at the standard level from May 2013 to May 2014? Among teachers who participated and those who did not participate in the specific types of professional activities prescribed by their evaluators, what percentage improved their summative ratings at the standard level from May 2013 to May 2014?

The first research question was addressed using the full population of 586 teachers who received prescriptions between May 2013 and February 2014 and who did not have an overall rating of unsatisfactory. The second research question was addressed using a random sample of these teachers stratified by age and racial/ethnic category, which consisted of a fourth of this population, or 148 teachers. The prescriptions for these teachers were coded for recommended activity types.

The remaining three research questions were addressed with the 248 teachers (42 percent) who completed a survey sent to all 586 teachers. There were no significant differences between respondents and nonrespondents in terms of the number of prescriptions received

---

### Box 3. Data and methods

The data consisted of teacher characteristics, evaluation ratings, and prescriptions for 586 teachers who:

- Received a rating of needs improvement or unsatisfactory in one or more standards on their 2013 summative or 2013/14 formative evaluation.
- Received a prescription between May 2013 and February 2014.<sup>1</sup>
- Did not have an overall rating of unsatisfactory on their 2013 summative evaluation.<sup>2</sup>

The district administered a survey to these teachers to capture the types of professional activities in which they participated during the evaluation cycle (May 2013 to May 2014) related to each of the four standards on the evaluation rubric. A total of 248 teachers (42 percent) completed the survey. Respondents differed from the full population of teachers in that respondents were more likely to be older, White, and female; otherwise, they were not significantly different from the nonrespondents regarding the number of prescriptions they received (see appendix C for more information about the representativeness of this sample). Interviews were conducted with six teachers and four principals about their experiences with the prescription process to provide context.

The following steps were taken to answer the research questions about each topic:

- **Professional needs identified in teachers' prescriptions.** The percentage of the full population of 586 teachers who received prescriptions in each standard was calculated.
- **Prescribed activities for professional practice and professional development activities.** The prescriptions for a random sample of a third of the teachers, stratified by age and race, were coded and analyzed to address the second research question. The activities in each prescription were coded into nine types (see table 1 and table D1 in appendix D) and frequencies were tabulated for each type (see table E2 in appendix E).
- **Teacher participation in professional practice and professional development activities.** Data from a district-administered teacher survey about professional development and practice activities were used to address the third research question. A nonresponse analysis was conducted. Frequencies and descriptive statistics were weighted to adjust for nonresponse.
- **Alignment of evaluators' prescriptions and teachers' actions.** The survey data were combined with the prescription data to assess whether teachers took the prescribed actions. The survey asked teachers to indicate whether they participated in professional activities associated with each standard and to identify the types of activities for each standard. The types of professional activities prescribed (such as coaching or taking courses) were examined for alignment to what teachers reported they did to address each standard. Alignment between prescriptions and reported activities is displayed in frequencies.
- **Improvement in evaluation ratings.** The percentage of teachers whose 2013/14 summative ratings were at least proficient were calculated by whether the teachers participated in any professional activities related to the standard in which they received a prescription and by the extent to which the types of reported activities for the standard aligned to what the evaluator prescribed.

#### Notes

1. The district determined that the time window for this analysis should be from May 2013, when teachers may have received prescriptions associated with their summative evaluation, through February 2014, when formative evaluations were due. Although teachers may receive prescriptions at any time during the year, ending the timeframe for receiving prescriptions in February left teachers time to take action on them by the May 2014 summative evaluation.

2. In 2013/14, 49 teachers had overall ratings of unsatisfactory. The district determined that the study population would not include these teachers because they were engaged in an intensive process of support and monitoring.

---

or the standards in which they were received. However, respondents were more likely to be older, White, and female. Because of these differences, the findings may not apply to all teachers who received prescriptions (see appendix C for more information about the survey sample). The number of teachers in the sample for each research question ranged from 148 to 586 (see table C2 in appendix C).

### **What the study found**

Teachers received prescriptions across all four standards. Prescriptions tended to include one or two professional activities. For all standards, teachers received more prescriptions with professional practice activities than professional development activities, but they reported participating in more professional development activities.

Teachers reported participating in more professional activities for standards 1 and 2 (the more instruction-based standards) than for standards 3 and 4 (family and community engagement and professional culture). For all standards, less than 40 percent of teachers who responded to the survey participated in all the activities their evaluators prescribed. However, at least 75 percent of teachers who received a prescription for standard 1 (curriculum, planning, and assessment) or standard 2 (teaching all students) and responded to the survey reported participating in at least one professional activity that addressed those standards.

A higher percentage of teachers who participated in activities of any type related to standard 1 received at least a proficient rating in this standard on their subsequent summative evaluation (64 percent) than did teachers with a prescription for standard 1 who did not participate in any activities related to the standard (38 percent). There was no statistically significant difference for other standards in the summative ratings of teachers who engaged in activities aligned to their prescriptions compared with those who did not.

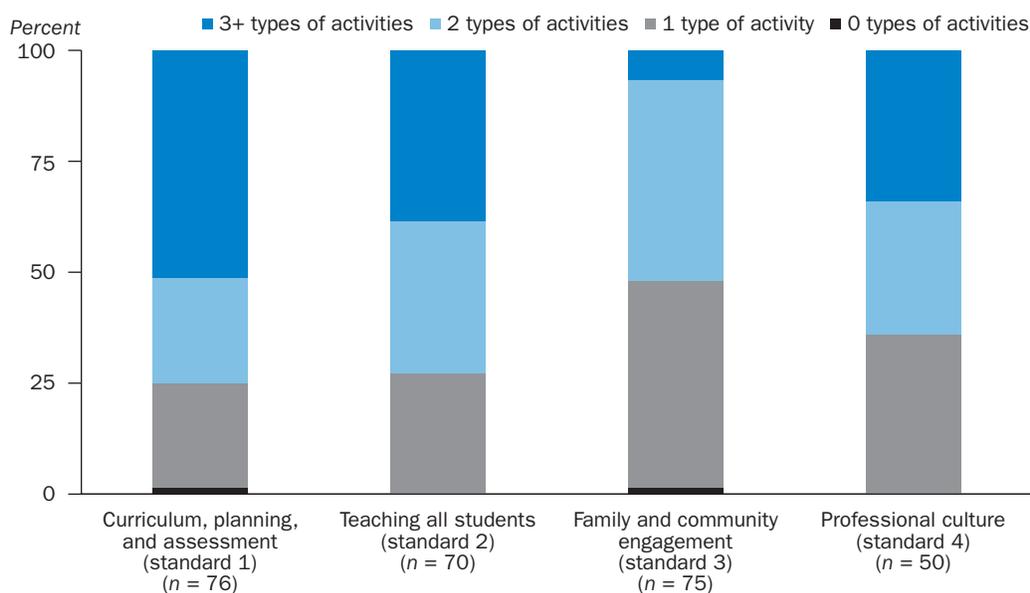
#### **Teachers received prescriptions across all four standards that tended to include one or two professional activities per prescription**

Forty-nine percent of teachers in the study population received a prescription for standard 1 (curriculum, planning, and assessment), 52 percent for standard 2 (teaching all students), 51 percent for standard 3 (family and community engagement), and 34 percent for standard 4 (professional culture).

Evaluators could prescribe whatever professional activities they deemed appropriate for a given standard, and the number of professional activities prescribed by standard varied. A majority of teachers in the sample received prescriptions with one or two types of professional activities for standards 2 (61 percent), 3 (92 percent), and 4 (66 percent), whereas standard 1 was the only standard for which the majority of teachers received prescriptions for three or more types of professional activities. By contrast, for standard 3, only 7 percent of teachers received prescriptions with three or more types of professional activities (figure 1). Teachers received the fewest prescriptions for standard 4 overall, and these prescriptions were evenly spread among one, two, and three or more types of activities. Few prescriptions included no activities (1 percent in standards 1 and 3).

**Forty-nine percent of teachers received a prescription for standard 1 (curriculum, planning, and assessment), 52 percent for standard 2 (teaching all students), 51 percent for standard 3 (family and community engagement), and 34 percent for standard 4 (professional culture)**

**Figure 1. Percentage of teachers in random sample who were prescribed one, two, or three or more types of activities, by standard, May 2013–February 2014**



**Note:** Based on a random sample of 148 teachers.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, and prescriptions.

*Evaluators tended not to prescribe professional development activities. Rather, they favored actions teachers could take on their own, either in the classroom or beyond it, that did not require them to work with other professionals, such as coaches, instructors, or evaluators*

**Teachers received more prescriptions with professional practice activities than with professional development activities for all standards**

Despite some variation across the standards, evaluators tended not to prescribe professional development activities. Rather, they favored actions teachers could take on their own, either in the classroom or beyond it, that did not require them to work with other professionals, such as coaches, instructors, or evaluators.

Nearly all teachers who received a prescription in each standard were prescribed professional practice activities for the standard (97–100 percent). In contrast, between 9 and 58 percent of teachers who received a prescription in each standard were prescribed professional development activities for the standard (table 2). Nearly all teachers (97 percent) who received a prescription under standard 1 (curriculum, planning, and assessment) were prescribed at least one professional practice activity whereas a much lower percentage (58 percent) received a prescription with at least one professional development activity. For standard 3 (family and community engagement), only 9 percent of teachers received a prescription that included a professional development activity.

Evaluators most frequently indicated that teachers should practice an instructional strategy (for standards 1 and 2) or other professional strategy (for standards 3 and 4; see table 2). For example, for standards 1 and 2, evaluators most frequently indicated that teachers should use instructional strategies, which might include the general direction to differentiate instruction or a specific direction to use a student engagement strategy such as “accountable talk” or “think, pair, share.” Teachers who received prescriptions for standards 3 and 4 most often were advised to use other professional strategies such as meeting deadlines, arriving at work on time, or following the expectations in the staff handbook.

**Table 2. Percentage of teachers in random sample prescribed each type of professional activity for each standard, May 2013–February 2014**

Type of professional activity	Curriculum, planning, and assessment (standard 1) (n = 76)	Teaching all students (standard 2) (n = 70)	Family and community engagement (standard 3) (n = 75)	Professional culture (standard 4) (n = 50)
Professional development	58	51	9	42
Workshop or course	17	13	0	18
Meeting with evaluator	12	13	3	4
Formal coaching or mentoring by a nonevaluator	18	16	1	4
Formal meeting with a colleague	38	30	8	28
Observation of a colleague	8	16	0	2
Professional practice	97	97	97	100
Document submission	59	26	55	42
Reading resources	18	19	1	6
Instructional strategies	87	93	4	28
Other professional strategies	13	14	87	76
None	1	0	1	0

**Note:** Teachers may have been prescribed more than one activity; therefore, the percentages may not sum to 100 percent. Based on a random sample of 148 teachers.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, and prescriptions.

*Some of the least frequently prescribed activities across all standards were professional development activities such as observing a colleague, meeting with an evaluator, or formal coaching or mentoring*

Document submission, such as submitting lesson plans to evaluators, was the second most frequent type of activity prescribed across three of the four standards.

Some of the least frequently prescribed activities across all standards were professional development activities such as observing a colleague, meeting with an evaluator, or formal coaching or mentoring.

Interviews with teachers and principals reinforced the finding that evaluators emphasized professional practice activities rather than professional development activities. Four of the six teachers interviewed noted that their prescriptions required submitting lesson plans. None of the interviewed teachers were advised to attend professional development courses.

I had to ... submit my ... lesson plans for all of my reading groups every Sunday night, and then [the evaluator] was to give me feedback on if she thought my lessons were going to move me out of the needs improvement category to proficient. (Teacher)

If [a teacher's] lessons weren't tight, weren't strong, or weren't focused, then one of the prescriptions might be that they had to turn in lesson plans ... weekly, to me. ... So I tried to base the prescription on what I thought were the areas of challenge for the teacher. (Principal)

The findings that follow are based, in part, on data from the district-administered survey in which 42 percent of teachers responded. Since there were some differences between the population of teachers who responded and did not respond to the survey, the findings may

not apply to all teachers who received prescriptions (see appendix C for more information about the survey sample).

**More teachers reported participating in professional activities, including both professional development and professional practice activities, for standards 1 and 2 than for standards 3 and 4**

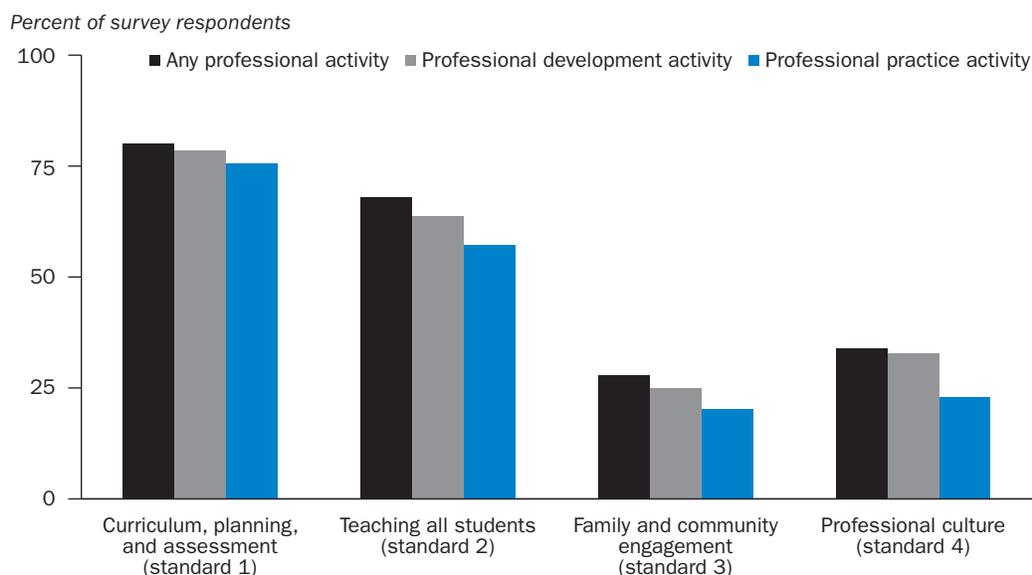
More teachers who responded to the survey (248 teachers, 42 percent of the population) reported participating in activities for standards 1 and 2 (the instruction-based standards) than for standard 3 (family and community engagement) and standard 4 (professional culture). Eighty percent reported participating in professional activities for standard 1, and 68 percent reported participating in professional activities for standard 2, while only 28 percent reported participating in any activity for standard 3 and 34 percent for standard 4 (figure 2).

For all standards, the percentage of teachers who reported participating in professional development and professional practice activities was similar. For example, for standard 1 (curriculum, planning, and assessment), 79 percent of teachers reported participating in professional development and 76 percent reported participating in professional practice activities (table 3).

The most commonly reported professional activity across all standards was participating in a workshop or course. For standards 1 and 2 the most commonly reported activities, in addition to workshop or course, were submitting documents (60 percent in standard

*For all standards, the percentage of teachers who reported participating in professional development and professional practice activities was similar*

**Figure 2. More survey respondents reported participating in any professional activities related to the instruction-based standards (standards 1 and 2) than professional activities related to the other two standards, May 2013–May 2014**



**Note:** Teachers may have participated in both professional development and professional practice activities. Reported percentages are weighted for nonresponse; 248 teachers were included in this analysis. See table E1 in appendix E for standard errors.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table 3. Percentage of survey respondents reporting participation in each type of professional activity for each standard, May 2013–May 2014**

Type of professional activity	Curriculum, planning, and assessment (standard 1)	Teaching all students (standard 2)	Family and community engagement (standard 3)	Professional culture (standard 4)
Professional development	79	64	25	33
Workshop or course	66	50	17	23
Meeting with evaluator	29	24	7	8
Formal coaching or mentoring by a nonevaluator	35	26	9	13
Formal meeting with a colleague	63	43	17	23
Observation of a colleague	41	29	6	10
Professional practice	76	57	20	23
Document submission	60	50	16	18
Reading resources	34	30	7	10
Specific strategies recommended by evaluator	49	39	15	15
None of the activities above	a	a	a	a
No activities	20	32	72	66

a. Not displayed due to small cell size.

**Note:** The survey did not distinguish between instructional strategies and other professional strategies; both are included in the category specific strategies recommended by evaluator. Reported percentages are weighted for nonresponse; 248 teachers completed the survey. Teachers may have participated in more than one activity; therefore, the percentages may not sum to 100 percent. See table E2 in appendix E for standard errors.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

1 and 50 percent in standard 2) followed by meeting with colleagues (63 in standard 1 and 43 percent in standard 2). Across all standards the least commonly reported activity was meeting with the evaluator. For each standard, between 20 percent and 72 percent of teachers reported participating in no activities.

While evaluators appear to have concentrated on professional practice activities that teachers could do independently, the interviewed teachers said they participated in many professional development activities, even more than evaluators prescribed. These professional activities included school or district workshops, team meetings, graduate courses, independent research, and work with teachers' union-based peer assistants. One teacher said she participated in a range of school-based professional development that was not aligned to her prescription but which she felt was valuable for her practice:

There was professional development. I wouldn't say that it was really related to the evaluation system at all. There was another administrator in the building who provided professional development. She'd have us meet in groups, and she found some really good articles to read, and we would read the article together, or come to a meeting having read the article, and discuss how it related to our practice, and try to make small, concrete changes to improve specific things. So there was professional development at the school but it wasn't tied to the evaluation system. (Teacher)

*While evaluators appear to have concentrated on professional practice activities that teachers could do independently, teachers said they participated in many professional development activities, even more than evaluators prescribed. These professional activities included school or district workshops, team meetings, graduate courses, independent research, and work with teachers' union-based peer assistants*

**For all standards, less than 40 percent of teachers participated in all the activities their evaluators prescribed**

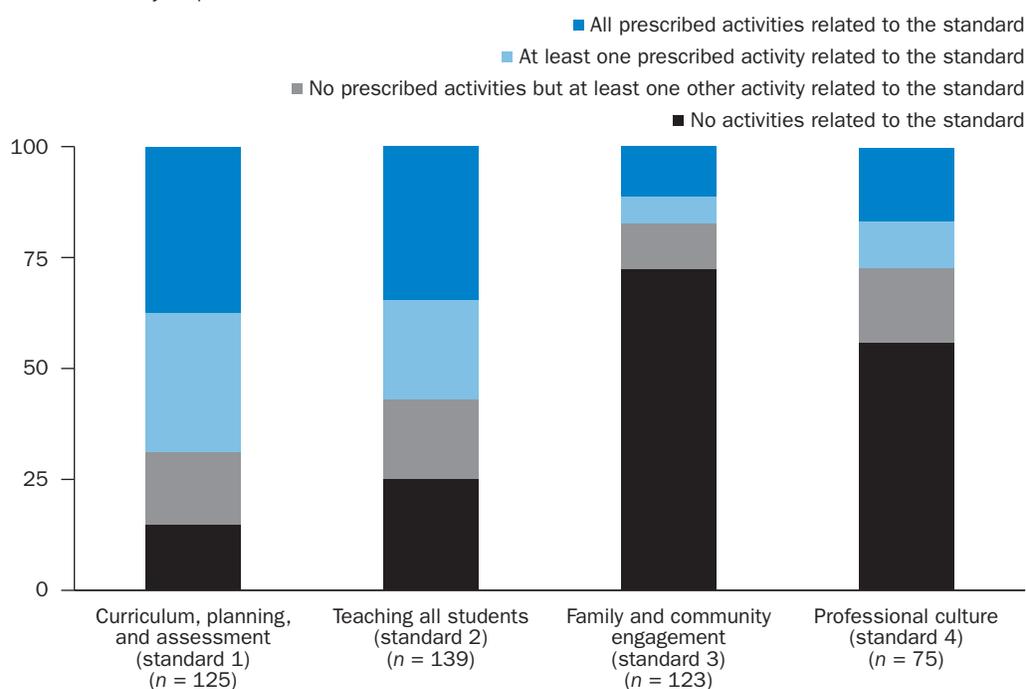
The majority of teachers responding to the survey reported participating in at least one professional activity prescribed by their evaluator for the two instruction-based standards: 69 percent for standard 1 (curriculum, planning, and assessment) and 58 percent for standard 2 (teaching all students) between May 2013 and May 2014 (figure 3). However, less than 40 percent of these teachers reported participating in all activity types prescribed by their evaluators. For example, while an evaluator might have prescribed participating in a course, meeting with colleagues, and submitting documents, the teacher might only have met with colleagues. In this example, the teacher’s professional activities only partially aligned with those prescribed by the evaluator. For standards 3 and 4, the misalignment between the professional activities evaluators prescribed and the activities the teachers reported was more pronounced: the majority of teachers participated in none of the activities prescribed by their evaluators (83 percent for standard 3, family and community engagement, and 73 percent for standard 4, professional culture).

*The majority of teachers reported participating in at least one professional activity prescribed by their evaluator for the two instruction-based standards: 69 percent for standard 1 (curriculum, planning, and assessment) and 58 percent for standard 2 (teaching all students)*

However, teachers did participate in other professional activities related to the standards. At least three-quarters of teachers who received a prescription for either standard 1 or standard 2 reported participating in at least one professional activity that addressed those standards (see table E3 in appendix E), whereas 28 percent of teachers who received a

**Figure 3. For all standards, less than 40 percent of survey respondents reported participating in all activities their evaluators prescribed, May 2013–May 2014**

Percent of survey respondents



**Note:** Reported percentages are weighted for nonresponse. Percentages may not sum to 100 because of rounding. See tables E3 and E4 in appendix E for standard errors.

**Source:** Authors’ analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, evaluator prescriptions, and a teacher survey.

prescription for standard 3 and 44 percent of teachers who received a prescription for standard 4 reported engaging in at least one professional activity related to those standards.

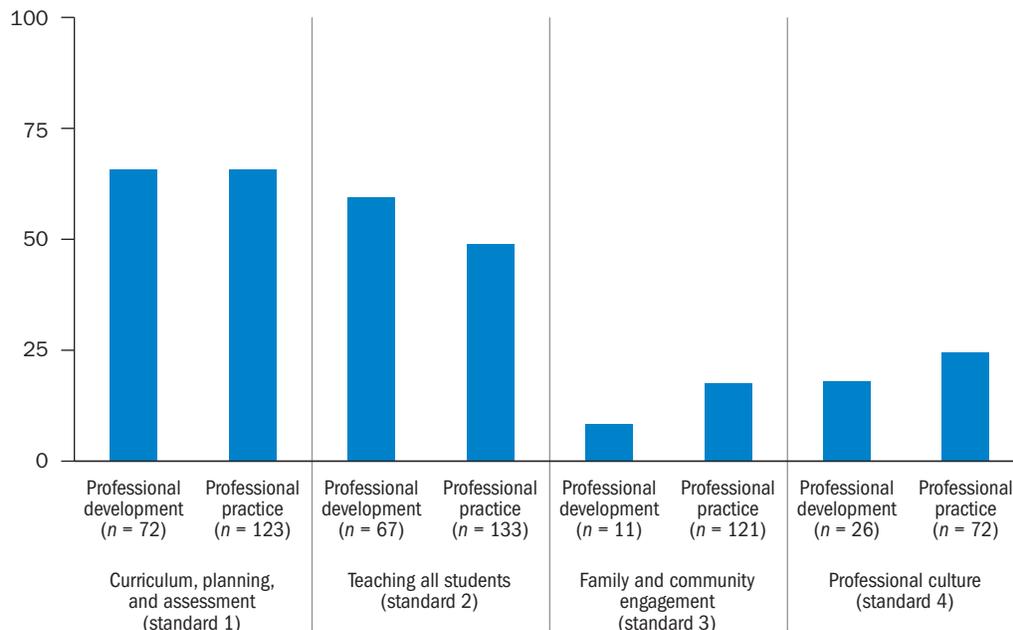
Evaluators prescribed professional development activities less frequently than they prescribed professional practice activities (see table 2). However, for standard 1 the percentage of teachers who reported participating in at least one professional development activity that their evaluators prescribed (66 percent) was the same as the percentage that reported participating in at least one professional practice activity prescribed by their evaluator (66 percent; figure 4). For standard 2 a higher percentage of teachers reported participating in at least one prescribed professional development activity (59 percent) than at least one prescribed professional practice activity (49 percent). The reverse was true for standards 3 and 4; a higher percentage of teachers reported participating in at least one prescribed professional practice activity.

When interview respondents were asked about the extent to which the prescriptions and their professional practice and development activities aligned, both teachers and principals

*For standard 1 the percentage of teachers who reported participating in at least one professional development activity that their evaluators prescribed was the same as the percentage that reported participating in at least one professional practice activity prescribed by their evaluator*

**Figure 4. For standard 2 a higher percentage of teachers reported participating in at least one prescribed professional development activity than at least one prescribed professional practice activity; the reverse was true for standards 3 and 4, May 2013–May 2014**

*Percent of survey respondents reporting participation in at least one prescribed activity*



**Note:** Reported percentages are weighted for nonresponse. The figure shows the percentage of teachers who participated in at least one prescribed activity for professional development or for professional practice, by standard. For example, of the teachers who received a prescription in standard 1, responded to the survey, and were prescribed professional development activities, 66 percent participated in at least one of the prescribed professional development activities. A teacher may have been prescribed both a professional development and professional practice activity in the same prescription. The number of teachers with prescriptions who responded to the survey was 125 for standard 1, 139 for standard 2, 123 for standard 3, and 75 for standard 4. See table E5 in appendix E for standard errors.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

indicated that the prescriptions and what teachers actually did were not deliberately or explicitly aligned. One teacher explained that her graduate course may have addressed student engagement, which was the content of her prescription, at some point, but she did not attend the course to deliberately address the content of her prescription nor did she see the course explicitly related to her prescription. She explained:

I don't think it was well aligned ... it was sort of accidental ... [the] course was extremely broad, but it wasn't ... specific to ... motivating students or student engagement strategies. That was part of the course but ... it wasn't like the whole means of the course. (Teacher)

**A higher percentage of teachers who participated in activities related to curriculum, planning, and assessment received at least a proficient rating in this standard on their subsequent summative evaluation than did teachers who did not participate**

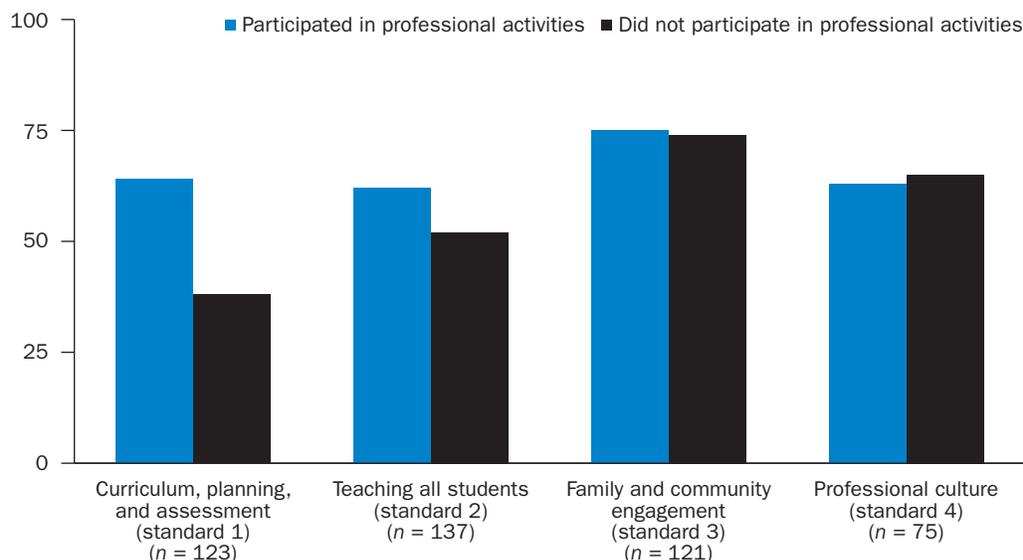
The percentage of teachers who were rated at least proficient on the summative 2013/14 rating for standard 1 (curriculum, planning, and assessment) was higher for those who participated in professional activities related to that standard than for those who did not. Of teachers with a prescription for standard 1 who participated in any professional activities related to that standard, 64 percent received a proficient or higher rating for standard 1 on their 2013/14 summative evaluation, compared with only 38 percent of teachers who were rated at least proficient in standard 1 but did not participate in any professional activity for standard 1, despite receiving a prescription for that standard (figure 5).

Standard 1 is the only standard with a statistically significant difference in summative 2013/14 ratings between teachers who did participate and those who did not participate in activities for the standard in which they received a prescription. Although 60 percent or more of teachers who received prescriptions in standards 2, 3, or 4 were rated at least proficient on their 2013/14 summative evaluation, there were no differences in the percentage rated proficient based on whether they participated in any professional activities in the standard for which they had a prescription (see figure 5 and table E7 in appendix E). Fulfilling the prescription is not required for a teacher's rating to improve; it is one of several factors an evaluator might consider in making rating decisions. The percentage of teachers who received at least a proficient rating on their summative evaluation did not vary by whether they participated in the particular types of activities their evaluator prescribed (see table E8 in appendix E).

***The percentage of teachers who were rated at least proficient on the summative 2013/14 rating for standard 1 (curriculum, planning, and assessment) was higher for those who participated in professional activities related to that standard than for those who did not***

**Figure 5. A significantly higher percentage of teachers who participated in professional activities related to standard 1 received at least a proficient rating on their 2013/14 summative evaluation than did teachers with a prescription for standard 1 who did not participate in any professional activities related to the standard**

Percent of survey respondents who were rated at least proficient on summative evaluation



*The findings suggest a tenuous connection between the educator evaluation system and professional learning*

**Note:** Percentages are weighted for nonresponse. The percentage of teachers who were rated at least proficient on the summative 2013/14 rating for standard 1 (curriculum, planning, and assessment) was significantly higher for those who participated in professional activities related to that standard than for those who did not. Standard 1 is the only standard with a statistically significant difference. See table E7 in appendix E for standard errors.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

### Implications of the study findings

District leadership viewed participation in this study as an opportunity to look closely at a system early in its implementation. Early results have already helped the district revise its support and training of evaluators. Other districts and states interested in aligning evaluation and professional development systems may benefit from the considerations and questions that emerged from this study. Although further work is needed to strengthen the connection between teacher evaluation and a comprehensive system of teacher support and development, this study takes the first step in illustrating the need for coherence between these related systems.

A key assumption of the district's prescription process was that evaluators would use the data gathered through multiple measures to provide substantive feedback in their prescriptions to teachers. In turn, teachers would take action based on the feedback to improve their practice and their summative evaluation rating. However, the findings suggest a tenuous connection between the educator evaluation system and professional learning.

This lack of alignment may have several possible explanations. First, there may have been a disparity between what evaluators believed teachers needed to do to improve their practice and what teachers could do or believed they needed to do. Second, teachers may not

have known how to implement their prescriptions, either because the prescriptions were too vague or because the activities prescribed were not available or the teachers did not know how to access them. Third, teachers may have lacked incentives for participating in the prescribed activities because they were under no direct contractual obligation to fulfill the specific terms of the prescription. Fourth, teachers may have been seeking more intensive professional development beyond that prescribed in the evaluations, whereas evaluators might have made professional practice suggestions tightly related to specific problems observed in the classroom to help teachers fulfill the requirements of the evaluation system. Further research is needed to understand why there was not greater alignment between the professional activities prescribed and the activities teachers reported related to each standard.

The finding that evaluators tended to prescribe professional practice activities over professional development activities might suggest that evaluators would benefit from more guidance on to how to write effective, actionable, and specific feedback that supports teacher improvement. The district provided training and guidance on the evaluation system, but evaluators may need more information about the range of professional activities available for teachers. Although it is not clear how evaluators decided what to write in their prescriptions, principal interview data suggest that evaluators need more training on how to use the prescription process to support teachers.

Finally, the finding that a higher percentage of teachers who participated in professional activities for standard 1 (curriculum, planning, and assessment) were at least proficient in that standard at the end of the evaluation cycle might be related to the relative importance of standard 1, which is focused on instructional practices, and may suggest that there is a more comprehensive menu of professional activities related to this area. Evaluators wrote prescriptions across all the standards, but this pattern appeared only for standard 1. Further research is needed to explore the range of professional activities available in the other standards.

The study raised several questions for further research:

- How do evaluators make decisions about which professional activities, including professional development, they recommend to teachers?
- What is the relationship between what evaluators prescribe and the resources in the district? Are there real or perceived limitations to what evaluators can recommend?
- Are some prescriptions more effective than others? Do certain recommendations yield better results in teachers' performance?
- What led to teachers' loose conformity to the prescriptions? Do teachers understand or disagree with them? Are the prescribed activities unclear or unrealistic?
- What are the most effective methods or approaches for training evaluators to write clear and actionable prescriptions?

### **Limitations of the study**

While the study raises important questions that may influence educator evaluation policy at the state and district levels, it is important to note its limitations. The most important limitation is that the findings should not be considered causal. The data are strictly descriptive and should not be interpreted to suggest that participation in a particular type

***That evaluators tended to prescribe professional practice activities over professional development activities might suggest that evaluators would benefit from more guidance on to how to write effective, actionable, and specific feedback that supports teacher improvement***

of activity or, in fact, participation in any professional activities will result in improvement in teacher ratings. The study was not designed to make claims about the effectiveness of any type of activity for improving teacher practice or ratings.

Also, the analysis did not distinguish between prescriptions that were general and those that provided significant detail and direction. For example, if an evaluator prescribed that a teacher differentiate instruction but did not provide detail about how to do this, the prescription was coded as an instructional strategy. If an evaluator detailed exactly how a teacher should differentiate instruction, indicating particular strategies and how to execute them in the classroom, the prescription was also coded as an instructional strategy. Thus, the analysis did not account for variation in the depth or specificity of the prescriptions, nor does the study make any claims about the quality of the prescriptions provided to teachers or suggest any relationship between quality of feedback and how teachers rated in their subsequent summative evaluations.

The coding of prescriptions may have inflated the percentage of teachers whose prescriptions and reported activities aligned. For example, if a prescription for standard 1 indicated that a teacher should participate in a particular workshop related to student engagement and the teacher reported participating in any workshop related to standard 1, but not necessarily the workshop the evaluator prescribed, the study team counted this as alignment between the prescription and the teacher's action. The survey did not probe for details such as the name of a workshop.

Further, only 42 percent of eligible teachers completed the survey. Those who responded did not represent all teachers with prescriptions. The survey respondents were more likely to be older, White, and female, compared with the full population of teachers with prescriptions. These respondents may have participated in more activities or different types of activities than nonrespondents. Further, the survey data about activities in which teachers engaged are self-reported. Whether teachers reported accurately what they did and whether interpretation of the types of activities provided in the survey matched the study teams' interpretation of the types indicated in the prescriptions cannot be known.

## **Appendix A. Extended literature review**

Studies about the U.S. policy context for new teacher evaluation systems and the connection between those systems and teachers' professional development activities are reviewed here. Few studies have examined the links between educator evaluation and professional development.

### **Policy context: Educator evaluation reform**

Many studies have called attention to the limitations of traditional educator evaluations to meaningfully differentiate among teachers' performance and relate teacher performance to student learning (Gordon, Kane, & Staiger, 2006; Heneman, Milanowski, Kimball, & Odden, 2006; Measures of Effective Teaching Project, 2012; Toch & Rothman, 2008; Weisberg, Sexton, Mulhern, & Keeling, 2009). Federal and state policymakers have taken a new interest in teacher evaluation in recent years. Federal grant competitions have encouraged or required recipients to reform teacher evaluation systems, tying them to teacher performance and including direct links to student achievement. For example, the U.S. Department of Education Race to the Top program, the School Improvement Grant program, and the Elementary and Secondary Education Act flexibility waiver applications include requirements that states reform their educator evaluation systems to include multiple measures, multiple rating scales, and evidence of student learning (Learning Point Associates, 2010; U.S. Department of Education, 2011, 2013). At the state level, by 2012, 23 states had enacted legislation to require annual evaluations of all teachers, and 43 states required annual evaluations of all new teachers (National Council on Teacher Quality, 2013). By 2013, 41 states required or recommended that these evaluations be conducted using multiple measures, such as combining student achievement results, student survey results, and classroom observations to determine a teacher's evaluation rating (Hull, 2013).

### **Limited research on new educator evaluation systems**

As redesigned teacher evaluation systems were adopted, studies began to examine these more rigorous approaches. Most research focused on the measures used in the evaluation (for example, standards and rubrics, student achievement) rather than the direct effects of evaluations on teacher performance. For example, researchers used value-added estimates to study the relationship between teachers' scores on observation rubrics and teachers' impact on student learning (see Ho & Kane, 2013 for a discussion of reliability in classroom observations and Kersting, Chen, & Stigler, 2013 for a discussion of value-added models). A number of studies have found that the use of multiple measures yields reliable evaluations, in which observational measures align with student performance measures (Sartain, Stoelinga, & Brown, 2011; Measures of Effective Teaching Project, 2010, 2012; Daley & Kim, 2010).

Another area of research has examined the overall reliability of evaluation systems for differentiating among teachers' performance (for example, Glazerman, Goldhaber, Loeb, Raudenbush, & Whitehurst, 2011). Because the new evaluation systems often include multiple rating scales (rather than the binary scales of previous systems), researchers have examined the distribution of teachers' ratings across these scales as a way to assess whether they effectively differentiate between good and bad teachers. For example, an Aspen Institute report about the Washington, DC, teacher evaluation system IMPACT, which

examined the trends in summative ratings before and after implementation of the new evaluation system, found that the new system elicited greater variation across the four-level rating scale, with fewer teachers performing at the highest level than in the previous system (Curtis, 2011). Thus the new evaluation system may provide greater differentiation of teacher quality, which would allow for easier identification of teachers who need additional professional development and support.

Few studies have directly examined the impact of evaluating teachers on their performance. In one exception, Taylor and Tyler (2011) studied the impact of an evaluation program in Cincinnati, Ohio, on midcareer teachers with more than five years of teaching experience. The evaluation was based on a teacher performance rubric, classroom observations, a review of teachers' work products, and feedback from the teachers during the school year, but not on student test scores. Teachers who participated in the evaluation program improved their effectiveness, as measured by increases in students' mathematics achievement test scores, during the year they were being evaluated and into subsequent years. However, similar gains were not found for reading achievement.

Finally, a recent examination of districts and charter schools that are part of an intensive Gates Foundation–supported initiative to enhance teacher effectiveness and student achievement found that teacher evaluation data were used to make decisions about professional development (Hamilton et al., 2014). All study sites had modified the ways in which they used evaluations to inform professional development decisions, including aligning professional development offerings to the evaluation data and providing individualized opportunities for teachers. However, teachers indicated that the districts could do more to effectively align the evaluations and professional offerings.

### **Understanding professional development**

The definitions, effectiveness, and the effective characteristics of professional development have been studied extensively.

Professional development has been defined broadly as “all activities that help education professionals develop the skills and knowledge required to achieve their school's education goals and meet the needs of students” (Chambers, Lam, & Mahitivanichcha, 2008, p. 4). Given this definition, it is often difficult to determine exactly which activities may be considered professional development. Professional development may range from trying out new strategies or skills for classroom instruction, to ongoing, job-embedded mentoring and coaching, to obtaining a certification through academic coursework (Great Schools Partnership, 2014).

Although there is limited statistically significant evidence of the impact of professional development on student outcomes (see Yoon, Duncan, Lee, Scarloss, & Shapley, 2007), the impact of professional development on teacher learning and practice has been a subject of investigation for several years. Reviews of the empirical literature, including a meta-analysis of 16 scientifically rigorous studies (Blank & de las Alas, 2009), have found that effective professional development can have a moderate effect on teachers' reported classroom practices and a small but significant effect on student achievement (Blank & de las Alas, 2009; Wallace, 2009; Yoon et al., 2007).

Findings from other studies, often based on teacher self-reporting and observation of classroom practice, have led to increased consensus about the core features of effective professional development (Supovitz, Mayer, & Kahle, 2000; Garet, Porter, Desimone, Birman, & Yoon, 2001; Weiss & Pasley, 2006; Penuel, Fishman, Yamaguchi, & Gallagher, 2007). Successful professional development is thought to be content-focused, ongoing, and coherent and to involve collective participation and active learning (Garet et al., 2001). Garet and others (2001) studied a national sample of teachers who participated in math and science professional development opportunities to investigate the impact of these core features based on teachers' self-report of increased knowledge and skills and changes in teacher practice. In a follow-up study, Penuel and others (2007) examined a more focused professional development initiative and found that changes in teachers' knowledge and practice, based on self-reporting, could be attributed to the presence of these features as well.

Professional development activities have been categorized as either traditional structured learning activities usually outside the classroom, or integrated activities that occur in the school on a regular basis and allow for more sustained learning (Chambers et al., 2008). There is an increasing consensus among researchers and practitioners that traditional professional development, which may include activities such as one-time workshops or off-site conferences, is insufficient to affect teachers' practice and support their ongoing professional growth. Rather, professional development that is job-embedded, content-focused, coherently linked to other instructional initiatives, and ongoing is widely considered to be the most effective way to improve teacher practice (Desimone, Porter, Garet, Yoon, & Birman, 2002; Guskey & Yoon, 2009; Odden, 2011; Behrstock-Sherratt & Jacques, 2012; Goe, Biggers, & Croft, 2012).

More recently, there has been a move toward expanding the traditional concept of professional development to include ongoing professional learning, defined as "planned and organized processes that actively engages educators in cycles of continuous improvement guided by the use of data and active inquiry around authentic problems and instructional practices" (Coggshall, 2012, p. 4). Under this perspective, there is less emphasis on single workshops or courses and more emphasis on job-embedded professional learning grounded in teachers' day-to-day classroom experiences (Croft, Coggshall, Dolan, Powers, & Killion, 2010). In a study of 30 elementary schools in a mid-size urban district, Parise and Spillane (2010) determined that both formal professional development and job-embedded professional learning were significantly and positively associated with teachers' reported changes in mathematics and English language arts instruction.

As a result of this research and other recommendations, the professional association Learning Forward offers a definition that is more consistent with job-embedded professional learning. Professional development is "a comprehensive, sustained, and intensive approach to improving teachers' and principals' effectiveness in raising student achievement" that includes the following characteristics: advances collective responsibility, is aligned with academic standards, is school-based and job-embedded, has a clear set of educator learning goals, and is regularly assessed for effectiveness related to improving teaching and student learning (Learning Forward, 2014).

To distinguish between professional development and professional practice, the study team drew on the definitions cited in this literature review. Specifically, Chambers and others' (2008) characterization of professional development as traditional (consisting of structured

learning activities that usually occur outside the classroom), or integrated (consisting of activities that occur in the school on a regular basis and allow for more sustained learning) guided the current study's description of professional development to include both traditional and integrated activities. Garet and others' (2001) core features of professional development—that it is content-focused, ongoing, coherent, and includes collective participation and active learning—guided the distinction between professional development, which involves interaction with mentors, colleagues, teachers, and evaluators, and professional practice, which is work the teacher undertakes independently. The concept of professional learning (Croft et al., 2010) was drawn on to define professional practice. Professional learning is grounded in day-to-day classroom experiences, and may include activities that go beyond either traditional or job-embedded professional development to include other practices in which a teacher may engage through day-to-day work at school. The current study places these day-to-day endeavors that a teacher conducts independent of colleagues, mentors, or peers in the professional practice category.

### **Aligning educator evaluation and professional development**

One critique of traditional evaluation systems is that they fail to provide teachers with recommendations for professional development or improvement (Weisberg et al., 2009). As a result, national organizations such as the Center on Great Teachers and Leaders increasingly recommend aligning educator evaluation and professional development systems to support improvements in instruction and student learning (for example, Coggshall et al., 2012; Curtis & Wiener, 2012; Danielson, 2012; Darling-Hammond & Stanford Center for Opportunity Policy in Education, 2012; Goe et al., 2014). A well-designed and well-implemented educator evaluation and professional development system would consist of teachers using common standards and metrics to understand and evaluate teaching practice, teachers using evidence-based feedback to reflect on and improve their performance, and schools integrating professional learning into collaborative cultures (Coggshall et al., 2012; Goe et al., 2012). Behrstock-Sherratt and Jacques (2012) argue that unless evaluation and professional support systems for teachers are linked, expectations for teachers are unclear, whereas in a standards-aligned professional development system, where evaluations and teacher professional development are based on the same set of standards, professional development helps teachers attain the standards assessed in the evaluation system. According to Goe and others (2012), if evaluation systems are to ensure that the value of the new standards, rubrics, and instruments is realized, high-quality aligned professional development is a necessary component. As of 2013, 31 states reported aligning their evaluation results with professional development opportunities for all teachers (Hull, 2013).

Despite this logic and these policy changes, there is scant research on the connection between evaluation and professional development; most such research comes from abroad. These studies affirm that timely, helpful feedback positively affects teachers' perceptions about the evaluation process and increases the likelihood that they will seek professional development based on the feedback. For example, Delvaux and others (2013) surveyed a representative sample of secondary school teachers in Flanders, Belgium, about their intent to participate in professional development based on the feedback from their evaluations. They found wide variation in teachers' reported professional development intentions, but teachers who were less experienced, who received useful feedback from their evaluators, and who believed that the school principal had a positive attitude were more likely to intend to engage in professional development following their evaluations (Delvaux et al.,

2013; see also Tuytens & Devos, 2011, 2014). This finding confirms research conducted in the United States that found teachers have more positive perceptions of evaluation systems when the evaluator provides timely feedback related to the teachers' content area and when the evaluator is perceived as helpful and reassuring (Milanowski & Heneman, 2001).

In short, teacher evaluation and professional development share the same purpose of helping teachers grow professionally, and therefore many researchers and practitioners argue that alignment between the two systems is essential to achieving this purpose. However, as Cogshall and others (2012) noted, there is little empirical evidence about the effectiveness of professional development aligned to an evaluation system for improving teacher practice.

## Appendix B. Sample prescription

This de-identified prescription was provided by an evaluator to a teacher between May 2013 and February 2014.

Standard 2. Teaching all students	
Indicator:	2A. Instruction
Problem statement:	Teacher does not consistently use instructional practices that are differentiated and likely to motivate and engage most students, including struggling learners and those with disabilities, in the lesson.
Evidence statement:	<ul style="list-style-type: none"><li>• Evidence 1: During each of my visits, Teacher used whole-class instruction for the mini lesson but did not regularly use cold calling strategies. Teacher also didn't use dip-sticking techniques during the mini lesson. As a result, teacher had insufficient information about the extent of student understanding and wasn't strategic about structuring groups or pairings during the independent work time. I did not observe any use of the techniques from The Skillful Teacher even though I recommended this as a resource.</li><li>• Evidence 2: During my visit on &lt;DATE&gt;, 33 percent of students were able to explain the objective, what they were learning, and why.</li><li>• Evidence 3: On &lt;DATE&gt;, I suggested that Teacher use a structured format for assigning partners and make sure students are clear about the content before asking them to work in pairs. Teacher did not follow through with using the strategy I suggested. When I observed on &lt;DATE&gt;, Teacher was still using informal groups or pairs and passed out a worksheet without making sure most students were clear about what had been covered in the mini lesson.</li></ul>
Prescription statement:	<ol style="list-style-type: none"><li>1. Teacher needs to use strategies to engage all students in the learning. He needs to conduct small group instruction during the independent work time. Teacher will use the various learning modalities of his students to engage them in the lesson. Teacher will use strategies such as "Turn and Talk" and "Think Pair Share" to encourage engagement.</li><li>2. Teacher needs to schedule time to work with the Literacy Coach during the week. In addition, Teacher should visit classrooms at the school to observe the Literacy Block.</li><li>3. Teacher should also enroll in the professional development literacy offered by the Literacy Department.</li></ol>

## **Appendix C. Data and methodology**

This appendix describes the study sample, data sources, and analyses strategies.

### **Sample**

The population for this study consisted of 586 teachers in an urban district in the Northeast & Islands Region who:

- Received a needs improvement or unsatisfactory rating in one or more standards on their 2013 summative or 2013/14 school year formative evaluation.
- Received a prescription between May 2013 and February 2014 (summative 2013 to formative 2014 evaluation period).
- Did not have an overall rating of unsatisfactory on their 2013 summative evaluation.

The district identified this group based on the following rationale. Teachers who had an overall rating of needs improvement, as well as teachers who had an overall rating of proficient but a less-than-proficient rating in at least one standard, represented the group that the district believed were most likely to make improvements in practice as a result of the prescription process and the most likely to be responsive to data collection efforts (for example, responsive to requests to complete the survey). This group also represented the largest group of teachers who received prescriptions. Teachers who received overall unsatisfactory ratings were placed on a different path for improvement.

The teachers in this study represented roughly 13 percent of the approximately 4,400 teachers employed by the district in 2012/13. Of the teachers with prescriptions, 61 percent were women. Prescriptions were not proportionally distributed by race: whereas 63 percent of the district's teachers were White, White teachers received only 48 percent of the prescriptions in the sample. Black teachers constituted 22 percent of the district's teacher population but received 34 percent of the prescriptions (table C1). The distribution of prescriptions was divided almost evenly among teachers in the three age groups; 34 percent of teachers were younger than 35, 36 percent were 35–50, and 30 percent were older than 50. Teachers at the middle school or grades K–8 level received the largest number of prescriptions (235 teachers or 40 percent). Elementary school teachers made up the smallest group of teachers with prescriptions, with 162 teachers (28 percent), and high school teachers were next, with 189 teachers (32 percent; see table C1).

The study posed the following research questions:

- In what standards did teachers receive prescriptions?
- For each standard, what professional activities did evaluators prescribe?
- What professional activities did teachers report they had participated in?
- Did the professional activities in which teachers participated align with the standards for which they received prescriptions and the activities that were prescribed?
- Among teachers who participated and those who did not participate in any professional activities in the standard in which they had a prescription, what percentage improved their summative ratings at the standard level from May 2013 to May 2014? Among teachers who participated and those who did not participate in the specific types of professional activities prescribed by their evaluators, what percentage improved their summative ratings at the standard level from May 2013 to May 2014?

**Table C1. Characteristics of teachers who received a prescription, May 2013–February 2014**

Characteristic	Number of teachers in study population	Percent of study population
<b>Gender</b>		
Female	359	61
Male	227	39
<b>Age</b>		
Younger than 35	200	34
35–50	212	36
Older than 50	174	30
<b>Race/ethnicity</b>		
White	282	48
Black	197	34
Other <sup>a</sup>	107	18
<b>School level taught</b>		
Elementary school	162	28
Middle school/K–8 <sup>b</sup>	235	40
High school	189	32
Total	586	100

**a.** Includes American Indian and Asian (this category was designated by the district and differs from Institute of Education Sciences’ racial/ethnic minority categories).

**b.** Includes schools that serve grades 6–8 and schools that serve grades K–8. The category is distinct from elementary schools, which serve grades K–5 generally.

**Source:** Authors’ analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

The first research question was addressed using the full population of teachers (586 teachers); the second research question was addressed using a random sample of 148 teachers—a fourth of the population—stratified by age and race (table C2). The remaining three research questions were addressed using the sample of teachers who completed the survey (248 teachers). Three teachers did not receive summative 2013/14 evaluation ratings and thus were not included in the sample for research question five. The analysis section describes how the samples compared with the full population of teachers. The unit of analysis for the study was the teacher, and all analyses were conducted separately for each of the four standards because the prescriptions and evaluation ratings (described below) were at the standards level.

The study team also conducted interviews with six teachers and four principals. A district leader sent an email to all teachers in the population (586 teachers) and all principals in the district to invite them to participate in an interview about the relationship between the evaluation system and professional development, particularly related to the prescription process. The district provided names and email addresses of 27 teachers randomly selected from respondents and all seven principals who agreed to an interview. The study team contacted teachers and principals to schedule phone interviews between June and July 2014. Six teachers and four principals agreed to participate in the role-specific interviews. The small sample for the interviews was likely due to the difficulty in making arrangements after school had closed for the summer. Since the sample was not representative of

**Table C2. Sample of teachers and number of prescriptions for each research question, May 2013–February 2014**

Research question	Sampling methodology	Number of teachers	Number of prescriptions
Research question 1	Population of teachers	586	1,090
Research question 2	Stratified random sample of teachers (strata are age and race)	148	271
Research question 3	Survey respondents	248	na
Research question 4	Survey respondents	248	462
Research question 5	Survey respondents who received a 2013/14 summative evaluation rating	245	456

na is not applicable because the research question does not analyze prescriptions.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

the full population of teachers, the interviews were used to provide context related to the prescription process and raise questions for further research.

## Data

**Prescription data.** The prescriptions indicate the standard on the professional practice rubric in which the teacher was expected to improve, a problem statement that outlined the issue identified, evidence to support the assignment of the rating, and a set of action steps required of the teacher. The standards in the evaluation rubric are:

- Standard 1: curriculum, planning, and assessment
- Standard 2: teaching all students
- Standard 3: family and community engagement
- Standard 4: professional culture

Initially the data included 1,486 prescriptions given between May 15, 2013, and February 1, 2014, for all 586 teachers in the target population. This time period was selected because it represented a full-cycle for the evaluation process. Prescriptions are typically given during the summative evaluation for 2012/13 and during formative evaluations, which were due by February 1. Prescriptions received after that due date were not included in the analysis because the district and the study team wanted to allow time for teachers to participate in the professional activities and improve their practice prior to the summative evaluation in May 2014.

Some teachers received more than one prescription. Note that each of the four standards may have had several indicators to which prescriptions could be written. Three scenarios in which a teacher may have received more than one prescription were:

- Prescriptions tagged to the same standard but to different indicators.
- Prescriptions tagged to the same standard and indicator but at different times during the evaluation cycle.
- Prescriptions tagged to different standards.

Since a teacher's evaluation rating was at the standard level (that is, evaluators rate teachers for each of the four standards), the first two scenarios were considered single prescriptions for the purpose of this analysis, whereas prescriptions in the third scenario, which

were tagged to different standards, were considered separate prescriptions. If a teacher received two prescriptions for standard 1 and each prescription included two activities, this was combined into a single prescription with four activities. However, if the teacher was prescribed the same activity for the same standard twice over the study timeframe, then that activity was not counted twice. The study team selected this approach so that prescriptions could be mapped to the teachers' ratings (there are no ratings at the indicator level) and also so that the unit of analysis for each standard was the teacher rather than the prescription. This was important because the teacher was the unit of analysis for the survey. After these rules were applied (that is, after combining prescriptions that were in the same standard for the same teacher), there were a total of 1,090 prescriptions.

**Evaluation ratings.** Teachers' 2013/14 summative evaluation ratings in each standard were used to determine whether they moved from less than proficient to at least proficient on a standard. The ratings are on a four-point scale: exemplary, proficient, needs improvement, and unsatisfactory. Teachers who received a rating of proficient or exemplary on their 2013/14 summative evaluation in May 2014 were considered to be at least proficient. Sixty percent of teachers were at least proficient on that evaluation in standard 1; 60 percent in standard 2; 74 percent in standard 3; and 64 percent in standard 4. See table E6 in appendix E.

**Teacher characteristics.** District administrative data on teacher characteristics, including age, race/ethnicity, gender, and level taught (elementary, middle, or high school) were provided for the full population of teachers. Age groups were younger than 35, ages 35–50, and older than 50. Race was categorized as White, Black, and other, which includes American Indian and Asian (this category was designated by the district and differs from Institute of Education Sciences' racial/ethnic minority categories).

**Survey data.** The district, in collaboration with the study team, developed a survey to collect information about the standards and the activity types addressed by the professional activities in which teachers engaged from May 2013–May 2014. The district administered the survey to all teachers in the target population (586 teachers). Combining the teachers' self-reported information from the survey with the prescription data allowed the study team to determine whether teachers participated in the professional activities prescribed by their evaluators. (See the analysis section below for more detail.) The district received survey responses from more than 300 teachers, but due to noncompletion by some respondents, the number of surveys matched to teachers' prescription data was 248. A copy of the survey is provided in appendix F.

**Interviews.** In interviews with teachers, study team members probed the needs identified in the prescriptions, the process teachers and principals employed to discuss and work with the prescriptions, and the types of professional activities in which teachers participated—both those related to their prescriptions and those that were in addition to the prescriptions. Teachers were also asked to comment on any professional activities that would have been helpful but were not available, as well as to consider what might contribute to a lack of relevant professional opportunities. The interviews with principals probed how they approached the prescription process including the preparation of prescriptions, the ongoing monitoring of teachers' progress, what training they received to support their use of the process, and their perceptions of the alignment of the prescriptions to available professional activities. The interviewer also asked principals to comment on whether teachers in their schools satisfactorily addressed the prescriptions and in the absence of

improvement, to what they might attribute teachers' lack of improvement. Finally, principals were asked to suggest changes to the professional development offerings as well as to the prescription process in general. Interviews were audio recorded and transcribed. Copies of the interview protocols are in appendix G.

## Analysis

The analyses conducted to address each research question are described below.

*In what standards did teachers receive prescriptions?* The percent of teachers in the study population who received a prescription for each standard was calculated.

*For each standard, what professional activities did evaluators prescribe?* A random sample of 148 teachers was selected from the population of 586 teachers because the analysis for research question 2 required qualitative coding of their prescription data. The sample was stratified by age and racial/ethnic category. The characteristics of the teachers in the sample and population were compared using a chi-square test (tables C3 and C4). There were no significant differences between the characteristics of the teachers included in the sample and the population of teachers, so the results can be considered representative of the study population.

**Table C3. Comparison between the population of teachers who received a prescription and the stratified random sample of 148 teachers, May 2013–February 2014**

Characteristic	Number of teachers in study population	Percent of study population	Number of teachers in study sample	Percent of study sample
<b>Gender</b>				
Female	359	61	88	59
Male	227	38	60	41
<b>Age</b>				
Younger than 35	200	34	51	34
35–50	212	36	54	36
Older than 50	174	29	43	29
<b>Race/ethnicity</b>				
White	282	48	71	48
Black	197	33	50	34
Other <sup>a</sup>	107	18	27	18
<b>School level taught</b>				
Elementary school	162	27	36	24
Middle school/K–8 <sup>b</sup>	235	40	52	35
High school	189	32	60	41
Total	586	100	148	100

**a.** Includes American Indian and Asian (this category was designated by the district and differs from Institute of Education Sciences' racial/ethnic minority categories).

**b.** Includes schools that serve grades 6–8 and schools that serve grades K–8. The category is distinct from elementary schools, which serve grades K–5 generally.

**Note:** The chi-square test statistic was 0.04 ( $p = .98$ ) for age, 0.3 ( $p = .6$ ) for female, 0.002 for race ( $p = .99$ ), and 1.3 for level taught ( $p = .52$ ).

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table C4. Teachers in the standard populations and in the random sample who received a prescription, by standard, May 2013–February 2014**

Standard	Number of teachers in study population	Percent of study population	Number of teachers in study sample	Percent of study sample
Curriculum, planning, and assessment (standard 1)	290	49	76	51
Teaching all students (standard 2)	307	52	70	47
Family and community engagement (standard 3)	296	51	75	51
Professional culture (standard 4)	197	34	50	34

**Note:** The chi-square test statistic was 1.04 ( $p = .79$ ). There were 586 teachers in the study population and 148 in the random sample.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

A total of 271 prescriptions were coded to address research question 2, and 462 were coded to address research questions 3 and 4. The content of the problem statement, the evidence statement, and the prescription action steps were used to determine the professional activity type or types and standard on the teacher rubric the prescription was meant to address. Professional activity codes were initially developed based on the literature on professional development (for example, Garet et al., 2001; Penuel et al., 2007; Supovitz, et al., 2000; Weiss & Pasley, 2006), and modified based on the input of district leadership about the types of professional activities in which teachers were likely to participate. Then the initial professional activity codes were applied to a sample of 15 prescriptions. Based on this initial coding, the study team identified additional codes needed to capture the range of activities prescribed. The additional codes were instructional strategies, other professional strategies, and document submission. Eleven codes were used to capture the professional activity types evaluators prescribed (nine activities plus none and missing data; see table D1 in appendix D). Of these, “none” (no professional activity prescribed) was employed in only two instances, and “missing data” was employed only twice when the prescription was cut off or incomplete.

Each prescription was then independently coded in Excel by two study team members according to the criteria in the coding dictionary (see table D1 in appendix D). Discrepancies in coding for professional activity types were reconciled through discussion, resulting in 100 percent agreement on all the final professional activity type codes.

The percentage of teachers whose evaluators prescribed each type of professional activity was calculated for each standard. The prescribed professional activities were grouped into professional development and professional practice. Professional development included workshops or courses, meetings with evaluator, formal coaching or mentoring by a non-evaluator, formal meeting with a colleague, and observation of a colleague. Professional practice included document submission, reading resources, instructional strategies, and other professional strategies.

The minimum, maximum, and mean number of types of professional activities prescribed to each teacher in a prescription were calculated. Finally, the percentage of teachers whose prescriptions included zero, one, two, and three or more types of professional activities was computed.

*What professional activities did teachers report they had participated in?* The district administered a survey to all teachers in the study population to determine whether they had participated in any professional activities in the standard for which they received a prescription, what type of professional activities they had participated in, and their general satisfaction with the professional activities. Although 304 teachers started the survey, 248 teachers (or 42 percent of the population) completed the necessary questions for this study. Only teachers who completed the survey were included in the analysis. A comparison of the age, gender, race/ethnicity, and grade level taught of teachers who completed the survey with those in the population using a chi-square test of percentages determined that older teachers and female teachers were more likely to respond to the survey (table C5).

A unit nonresponse bias analysis was conducted using a logit model where the dependent variable indicated response status (1 = response, 0 = nonresponse) and the independent variables were age as a continuous variable; race categorized as Black, White (reference category), and other; and female. The significant predictors were age, Black, and female. In addition, the study team conducted nonresponse analysis to address any differences in respondents and nonrespondents based on the standard in which teachers received prescriptions, the number of prescriptions, the number of standards in which they received prescriptions, and the summative 2012/13 overall and standard ratings (table C6). The only statistically significant difference between respondents and nonrespondents was for the ratings for standard 4. However, when the 2012/13 standard 4 rating was included in

**Table C5. Characteristics of teachers who received a prescription, May 2013–February 2014, who responded and did not respond to the survey**

Characteristics	Respondents		Nonrespondents	
	Number	Percent	Number	Percent
<b>Gender</b>				
Female	164	66.1	195	57.7
Male	84	33.9	143	43.3
<b>Age</b>				
Younger than 35	72	29.0	128	37.9
35–50	95	38.3	117	34.6
Older than 50	81	32.7	93	27.5
<b>Race/ethnicity</b>				
White	130	52.4	152	45.0
Black	75	30.2	122	36.1
Other <sup>a</sup>	43	17.3	64	18.9
<b>School level taught</b>				
Elementary school	68	27.4	94	27.8
Middle school/K–8 <sup>b</sup>	99	39.9	136	40.2
High school	81	32.7	108	32.0
Total	248		338	

**a.** Includes American Indian and Asian (this category was designated by the district and differs from Institute of Education Sciences' racial/ethnic minority categories).

**b.** Includes schools that serve grades 6–8 and schools that serve grades K–8. The category is distinct from elementary schools, which serve grades K–5 generally.

**Note:** The chi-square test statistic was 5.09 ( $p = .079$ ) for age, 4.3 ( $p = .038$ ) for female, 3.3 for race ( $p = .191$ ), and 0.03 ( $p = .98$ ) for level taught.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table C6. Teachers' evaluation ratings and number of prescriptions for teachers who responded and did not respond to the survey, May 2013–February 2014**

Ratings by standard and overall and number of prescriptions received	Respondents		Nonrespondents	
	Number	Percent	Number	Percent
<b>Curriculum, planning, and assessment (standard 1) 2012/13 rating</b>				
Exemplary	4	2	6	2
Proficient	150	60	198	59
Needs improvement or unsatisfactory	43	17	63	19
No summative 2012/13 rating <sup>a</sup>	51	21	71	21
<b>Teaching all students (standard 2) 2012/13 rating</b>				
Proficient or exemplary	145	59	192	57
Needs improvement	52	21	70	21
Unsatisfactory	0	0	5	1
No summative 2012/13 rating <sup>a</sup>	51	21	71	21
<b>Family and community engagement (standard 3) 2012/13 rating</b>				
Exemplary	4	2	7	2
Proficient	146	59	193	57
Needs improvement or unsatisfactory	47	19	67	20
No summative 2012/13 rating <sup>a</sup>	51	21	71	21
<b>Professional culture (standard 4) 2012/13 rating</b>				
Exemplary	5	2	12	4
Proficient	165	67	190	56
Needs improvement or unsatisfactory	28	11	61	19
No summative 2012/13 rating <sup>a</sup>	51	21	71	21
<b>Overall 2012/13 rating</b>				
Proficient or exemplary	133	54	179	53
Needs improvement	64	26	90	27
No summative 2012/13 rating <sup>a</sup>	51	21	69	20
<b>Number of standards in which teachers received prescriptions</b>				
1	15	4	28	8
2	161	48	198	59
3	40	12	71	21
4	32	9	41	12
<b>Number of prescriptions</b>				
1	100	30	148	44
2	63	19	72	21
3	34	10	45	13
4 or more	51	15	73	22
Mean number of prescriptions	2.5		2.6	
<b>Standards in which teachers received prescriptions</b>				
Curriculum, planning, and assessment (standard 1)	125	50	165	49
Teaching all students (standard 2)	139	56	168	50
Family and community engagement (standard 3)	123	50	173	51
Professional culture (standard 4)	75	30	122	36
Number of teachers	248		338	

(continued)

---

**Table C6. Teachers' evaluation ratings and number of prescriptions for teachers who responded and did not respond to the survey, May 2013–February 2014**

*(continued)*

**Note:** In some cases the categories were collapsed because of the small number of teachers who received a particular rating. The chi-square test statistic was 0.94 ( $p = .92$ ) for standard 1 2012/13 rating, 1.88 ( $p = .39$ ) for standard 2 2012/13 rating, 1.72 ( $p = .79$ ) for standard 3 2012/13 rating, 10.22 ( $p = .04$ ) for standard 4 2012/13 rating, 3.78 ( $p = .29$ ) for the number of standards in which a teacher received prescriptions, and 1.89 ( $p = .39$ ) for the total number of prescriptions; and the percentage of teachers who received a prescription was 0.14 ( $p = .70$ ) for standard 1, 2.31 ( $p = .13$ ) for standard 2, 0.14 ( $p = .70$ ) for standard 3, and 2.20 ( $p = .14$ ) for standard 4. The  $t$  statistic comparing the mean number of prescriptions between respondents and nonrespondents was 0.25 ( $p = .80$ ).

**a.** While these teachers did not receive a summative 2012/13 rating, each prescription is accompanied by a rating, which must be either unsatisfactory or needs improvement; all teachers received a rating for each prescription that they received between May 2013 and February 2014.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

---

the nonresponse regression as an independent variable with the previously specified characteristics of age, gender, and race/ethnicity, it was not significant.

The study team then calculated nonresponse weights equal to the inverse of the estimated probability of response from the model that included age, race/ethnicity, and gender. The remaining research questions incorporate the nonresponse weights when calculating the descriptive statistics.

The weighted percentage of teachers who reported having participated in professional activities in each standard and the types of professional activities (such as online courses and coaching) in which the teachers engaged was calculated. In some cases, the response options listed in the survey for the types of professional activities were more detailed than the codes used to analyze the prescriptions. Thus, some activity types in the survey were combined to align with the prescriptions' professional activity codes. Specifically, document submission includes the survey response options document submission with feedback and document submission without feedback; and courses included the following response options: a school- or district-based workshop during the academic year, a summer workshop, a face-to-face course located outside of your district, an online course or webinar, and a conference. (See appendix F for a copy of the survey.)

As described earlier, interviews with teachers and principals provided contextual information about the needs identified in a small group of teachers' prescriptions, as well as details about the process in which teachers and principals engaged to communicate and address the prescriptions. Ten interviews were recorded and transcribed. Initially, the study team employed a set of a priori codes aligned with the questions in the interview protocols (appendix G). Two members of the study team coded two common interviews employing the initial codes and determined that, in addition to the protocol questions to guide coding, the following relevant codes were necessary: school culture, relationship evaluator/teacher, and evaluation system. Once the final codes were determined, the two study team members coded all interviews. Large excerpts of interview data were grouped within these codes, and the study team read the full coded interviews to identify common themes across the interviews. These common themes were employed to confirm findings from other data sources and to raise additional questions warranting further research related to the alignment of professional development and evaluation in a large urban district.

*Did the professional activities in which teachers participated align with the standards for which they received prescriptions and the activities that were prescribed?*

The term alignment is used to refer to the extent to which the professional activities in which teachers reported participating were related to the standard in which they received prescriptions and the professional activities prescribed. The survey data were used to determine if a teacher participated in professional activities for the standard in which the teacher received a prescription. The report includes nonresponse weighted proportions of teachers who participated in any professional activity in the standard tagged to their prescription.

To determine whether teachers engaged in the type of professional activity (for example, courses or coaching) prescribed by the evaluator, the study team examined the alignment between the types of activities identified in a given prescription and what a teacher reported to have done related to that standard. If none of the types of professional activities that the teacher reported participating in matched the types of professional activities prescribed, then the teacher was categorized as having participated in none of the activities prescribed. If at least one of the professional activities the teacher reported participating in for the standard matched at least one of the professional activities prescribed, but the teacher did not participate in all the activities prescribed, then the teacher was categorized as having participated in at least one prescribed activity, but not all. For example, a teacher may have reported participating in an online course to address a prescription under standard 1 but did not report participating in weekly meetings with a supervisor, which was another type of professional activity the evaluator prescribed for standard 1. Therefore, the teacher participated in at least one activity prescribed, but not all. Finally, a teacher who reported engaging in all of the activities prescribed was categorized as engaging in all activities prescribed.

If the evaluator prescribed other professional strategies or other instructional strategies and the teacher reported that he or she tried specific strategies recommended by the evaluator, the study team considered this as a match between the prescription and the teacher's action. The study team calculated the nonresponse weighted proportion of teachers who participated in none of the types of professional activities prescribed, at least one type prescribed, but not all, and all types prescribed.

*Among teachers who participated and those who did not participate in any professional activities in the standard in which they had a prescription, what percentage improved their summative ratings at the standard level from May 2013 to May 2014? Among teachers who participated and those who did not participate in the specific types of professional activities prescribed by their evaluators, what percentage improved their summative ratings at the standard level from May 2013 to May 2014?* The study team created a variable equal to 1 to indicate a teacher who received a summative evaluation rating in May 2014 of at least proficient (a rating of proficient or exemplary) on the standards tied to the teacher's prescriptions, and 0 otherwise. Since these teachers previously had a rating of needs improvement or unsatisfactory in the standard tied to the prescription, a variable equal to 1 indicated that the teacher's rating improved.

The study team then calculated the percentage of teachers whose summative ratings were at least proficient by whether the teacher participated in any professional activity in the standard tagged to the prescription and by the extent to which the types of professional

activities in which the teacher participated aligned to what was prescribed by the evaluator. The nonresponse-weighted percentages of teachers rated at least proficient on their summative evaluations by these measures of alignment are reported in tables E7 and E8 in appendix E. Wald tests were used to determine whether the differences in the percentages rated at least proficient were significant.

## Appendix D. Coding dictionary

A coding dictionary was developed to analyze the prescription narratives. Coding allowed the study team to summarize and quantify the professional activities in the narratives and to match the activities prescribed to the activities reported in the survey to study the alignment between what evaluators prescribed and what teachers reported.

**Table D1. Coding dictionary**

Professional activities type code	Definition
Professional development	
Workshop or course	Any reference to a workshop, course, class, or professional development offering. If the prescription suggests that the teacher “take PD [professional development],” “seek out PD,” or “find PD,” the prescription would be coded as workshop/course. If the prescription says, “avail yourself” of PD, do NOT code.
Meeting with evaluator	Any reference to meeting with the evaluator or supervisor; in instances where the prescription requests the teacher meet with “me,” the prescription would code this as meeting with evaluator.
Formal coaching or mentoring by a nonevaluator	Any reference to a “coach” or “coaching” and reference to meeting, seeing, scheduling, or talking with a “coach” would be coded here. If the evidence or problem statement mentions the coach but does not say to meet with the coach in the future, do not code as “formal coaching.” If the coach or mentor is mentioned as a future relationship or ongoing relationship, then code; but if the coach is mentioned as a past relationship, this is not formal coaching.
Formal meeting with a colleague	Meetings with colleagues may include references to meeting with content specialists, peer assistants, other teachers, collaboration during common planning time, or, in general, being assigned the task of collaborating on a common product such as a lesson, unit, communication plan, and the like. The language must be stronger than “avail yourself” or “take advantage” of colleagues, and the like, in order to be coded as a meeting (either with an evaluator or with colleagues). In some cases the language may indicate “communicate with.” In such a case the prescription would be coded as formal meeting with a colleague if that communication appears to be required of the teacher. “Colleagues” can include new teacher developer, other administrators not the evaluator.
Observation of a colleague	Any reference to visiting another teacher’s classroom, sitting in on another class, and the like.
Professional practice	
Document submission	Any reference to submitting or providing evidence, materials, lesson plans, time logs, etc. These submissions do not have to be regular or reviewed, though they might be. A recommendation to create a document, for example, “maintain a log,” is not document submission. If the evaluator requires that such a document be reviewed or discussed, it is document submission.
Reading resources	Any suggestion to read a book, a chapter, an article, a website, specific guidance materials, videos, and the like would be coded as reading resources. Any reference to “literature,” even if not specified, would qualify as reading resource. Note: referring the teacher to review the teacher rubric or the state standards does NOT count as “reading resources” and should NOT be coded as such unless it refers to examples or resources about the rubric. Can include reading students’ documents such as individualized education programs for student in special education. Just a reference to “research” is not reading resource. Lesson plan templates are reading resources.
Instructional strategies	Any strategies that have to do with what the teacher does in the classroom with students, including lesson planning, lesson execution, assessment, and classroom management, including student behavior that is addressed by the teacher in classroom. Even reference to a general idea for implementation, such as “differentiate instruction” would be coded, even though there’s no further explanation in the prescription.

*(continued)*

---

**Table D1. Coding dictionary** *(continued)*

Professional activities type code	Definition
Other professional strategies	Any strategies suggested that have to do with how the teacher interacts with other adults in the building, with parents, and with the community. Such strategies also include any reference to meeting deadlines, punctuality, completing paperwork, and attendance. Even reference to a general idea for implementation, such as “contact with parents, including logs” would be coded, even though there’s no further explanation in the prescription. (Can include recommendations that teachers suggest how parents can work with students at home.) Any reference to implementing an “action plan.”
None	If the prescription includes no reference to any of the above type codes, code as none.
Missing data	If the prescription appears to be cut off, code the full prescription as missing data in the additional column at the end. If the prescription refers to a document that is not included and it is impossible to make a determination of the type, it constitutes missing data. If the prescription is specific about type, and there is no document for content, defer to the evaluator.

---

## Appendix E. Supplementary tables

This appendix includes the percentages and standard errors that align to the figures and tables in the main text.

**Table E1. Percentage of survey respondents who reported participating or not participating in professional activities related to each standard, May 2013–May 2014**

Standard	Participated		Did not participate	
	Percent	Standard error	Percent	Standard error
Curriculum, planning, and assessment (standard 1) ( <i>n</i> = 201)	80	0.03	20	0.03
Teaching all students (standard 2) ( <i>n</i> = 172)	68	0.03	32	0.03
Family and community engagement (standard 3) ( <i>n</i> = 72)	28	0.03	72	0.03
Professional culture (standard 4) ( <i>n</i> = 87)	34	0.03	66	0.03

**Note:** Reported percentages are weighted for nonresponse. *n* = 248 teachers.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table E2. Percentage of survey respondents who reported participating in each type of professional activity, by standard, May 2013–May 2014**

Professional activity	Curriculum, planning, and assessment (standard 1) ( <i>n</i> = 201)		Teaching all students (standard 2) ( <i>n</i> = 172)		Family and community engagement (standard 3) ( <i>n</i> = 72)		Professional culture (standard 4) ( <i>n</i> = 87)	
	Percent	Standard error	Percent	Standard error	Percent	Standard error	Percent	Standard error
Professional development	79	0.03	64	0.03	25	0.03	33	0.03
Workshop or course	66	0.03	50	0.03	17	0.03	23	0.03
Meeting with evaluator	29	0.03	24	0.03	7	0.02	8	0.02
Formal coaching or mentoring by a nonevaluator	35	0.03	26	0.03	9	0.02	13	0.02
Formal meeting with a colleague	63	0.03	43	0.03	17	0.01	23	0.03
Observation of a colleague	41	0.03	29	0.03	6	0.02	10	0.02
Professional practice	75	0.03	57	0.03	20	0.03	67	0.03
Document submission	60	0.03	50	0.03	16	0.02	18	0.02
Reading resources	34	0.03	30	0.03	7	0.02	10	0.02
Specific strategies recommended by evaluator	49	0.03	39	0.03	15	0.02	15	0.02
None of the activities above	..	..	..	..	..	..	0	na
No activities	20	0.03	32	0.03	72	0.03	66	0.03

na is not applicable.

.. the value is negligible and is not displayed.

**Note:** Reported percentages are weighted for nonresponse. *n* = 248 teachers.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table E3. Percentage of survey respondents who reported participating in any professional activities in the standard in which they received a prescription, May 2013–May 2014**

Standard	Percent	Standard error
Curriculum, planning, and assessment (standard 1) (n = 125)	85	0.03
Teaching all students (standard 2) (n = 139)	75	0.04
Family and community engagement (standard 3) (n = 123)	28	0.04
Professional culture (standard 4) (n = 75)	44	0.06

**Note:** Reported percentages are weighted for nonresponse.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table E4. Percentage of survey respondents who reported participating in the type of professional activities prescribed by their evaluators, by standard, May 2013–May 2014**

Number of professional activities	Curriculum, planning, and assessment (standard 1) (n = 125)		Teaching all students (standard 2) (n = 139)		Family and community engagement (standard 3) (n = 123)		Professional culture (standard 4) (n = 75)	
	Percent	Standard error	Percent	Standard error	Percent	Standard error	Percent	Standard error
No activities in the standard	15	0.03	25	0.04	72	0.04	56	0.06
None of the activities prescribed	16	0.03	18	0.03	10	0.03	17	0.04
No prescribed activities but at least one other activity related to the standard	31	0.04	23	0.04	6	0.02	11	0.03
All activities prescribed	38	0.04	35	0.04	11	0.03	17	0.04

**Note:** Reported percentages are weighted for nonresponse.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table E5. Percentage of survey respondents who reported participating in at least one professional activity prescribed by their evaluators, by standard, May 2013–May 2014**

Standard	Professional development			Professional practice		
	Number prescribed activity	Percent engaged in activity	Standard error	Number prescribed activity	Percent engaged in activity	Standard error
Curriculum, planning, and assessment (standard 1)	72	65.9	0.06	123	65.7	0.04
Teaching all students (standard 2)	67	59.4	0.06	133	49.1	0.04
Family and community engagement (standard 3)	11	0.09	0.09	121	17.6	0.03
Professional culture (standard 4)	26	18.1	0.08	72	24.5	0.05

**Note:** Reported percentages are weighted for nonresponse. The number of teachers with prescriptions who responded to the survey was 125 for standard 1, 139 for standard 2, 123 for standard 3, and 75 for standard 4.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table E6. Percentage of survey respondents who received a prescription, May 2013–February 2014, and were rated at least proficient on the 2013/14 summative evaluation, by standard**

Standard	Percent	Standard error
Curriculum, planning, and assessment (standard 1) (n = 123)	60	0.04
Teaching all students (standard 2) (n = 137)	60	0.04
Family and community engagement (standard 3) (n = 121)	74	0.04
Professional culture (standard 4) (n = 75)	64	0.06

**Note:** Reported percentages are weighted for nonresponse.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, a teacher survey, and interviews.

**Table E7. Percentage of survey respondents who were rated at least proficient on their 2013/14 summative evaluation by whether they participated in any professional activities related to the standard for which they had a prescription, by standard, May 2013–May 2014**

Participation in professional activity	Curriculum, planning, and assessment (standard 1) (n = 123)		Teaching all students (standard 2) (n = 137)		Family and community engagement (standard 3) (n = 121)		Professional culture (standard 4) (n = 75)	
	Percent	Standard error	Percent	Standard error	Percent	Standard error	Percent	Standard error
Did not participate in professional activities	38	0.12	52	0.09	74	0.05	65	0.08
Participated in professional activities	64	0.05	62	0.05	75	0.07	63	0.09
Wald test	F(1, 122) = 4.17*		F(1, 136) = 0.88		F(1, 120) = 0.01		F(1, 74) = 0.03	

\* Significant at  $p < .05$ .

**Note:** Reported percentages are weighted for nonresponse.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, and a teacher survey.

**Table E8. Percentage of survey respondents who were rated at least proficient on their 2013/14 summative evaluation by whether they participated in the prescribed types of professional activities in the standard for which they had a prescription, by standard, May 2013–May 2014**

Type of prescribed professional activities engaged in	Curriculum, planning, and assessment (standard 1) (n = 123)		Teaching all students (standard 2) (n = 137)		Family and community engagement (standard 3) (n = 121)		Professional culture (standard 4) (n = 75)	
	Percent	Standard error	Percent	Standard error	Percent	Standard error	Percent	Standard error
None of the types prescribed	51	0.08	58	0.07	75	0.04	67	0.06
At least one type prescribed, but not all	69	0.08	56	0.09	78	0.14	67	0.16
All types prescribed	61	0.07	64	0.07	61	0.14	53	0.15
Wald test	F(2, 122) = 1.21		F(2, 136) = 0.34		F(2, 120) = 0.49		F(2, 74) = 0.38	

**Note:** Reported percentages are weighted for nonresponse. There were no significant differences.

**Source:** Authors' analysis based on 2013 and 2014 district data on teacher characteristics and evaluation ratings, prescriptions, a teacher survey.

## Appendix F. Survey items

This appendix includes the full text of the survey administered to all teachers who received a needs improvement rating in one or more standards, and therefore had at least one prescription.

Dear Educator, [Name of District] is collaborating with the Regional Educational Laboratory Northeast & Islands (REL-NEI) to study professional development in the district's new evaluation system. Specifically, this survey is about the types of professional development that teachers engaged in between May 2013 and May 2014 and the standards and indicators that the professional development addressed. In addition, there are questions about the type and duration of the professional development activities and about school climate. For the purposes of this survey, professional development is a broad term that includes both formal and informal activities, feedback, interactions, and efforts. Your participation in this survey is voluntary. You may choose to skip any question(s). Please know that your responses to this survey are confidential and will be used for research purposes only.

1. Did you engage in any professional development or support related to Standard 1: Curriculum, Planning, and Assessment?
  - a. Yes
  - b. No
  
2. For each type of professional development or support listed related to Standard 1: Curriculum, Planning, and Assessment, indicate whether you engaged in that type of professional development and indicate its duration.
  - a. Did you engage in this type of professional development? (Yes/No)
    - i. Tried specific strategies recommended by evaluator
    - ii. Read resources recommended by evaluator
    - iii. Regular meetings with evaluator
    - iv. Document (e.g., lesson plans, data, artifacts) submission with feedback
    - v. Document (e.g., lesson plans, data, artifacts) submission without feedback
    - vi. Observations of a colleague
    - vii. Formal meetings with colleagues (e.g., data team meetings, professional learning community)
    - viii. Formal coaching or mentoring by non-evaluator
    - ix. A school- or district-based workshop during the academic year
    - x. A summer workshop
    - xi. A face-to-face course located outside of your district
    - xii. An online course or webinar
    - xiii. Attended a conference
  - b. What was the duration?
    - i. One day or less
    - ii. Two or more consecutive days
    - iii. Two or more days spread over several weeks
    - iv. Two or more days spread over several months (e.g., a semester)
    - v. Ongoing (no definitive end date)

3. Which of the following indicators did the professional development address? (select all that apply)
  - a. Indicator A — Curriculum & Planning
  - b. Indicator B — Assessment
  - c. Indicator C — Analysis
  - d. I don't know
  
4. Did you engage in any professional development or support related to Standard 2: Teaching All Students?
  - a. Yes
  - b. No
  
5. For each type of professional development or support listed related to Standard 2: Teaching All Students, indicate whether you engaged in that type of professional development and indicate its duration.
  - a. Did you engage in this type of professional development? (Yes/No)
    - i. Tried specific strategies recommended by evaluator
    - ii. Read resources recommended by evaluator
    - iii. Regular meetings with evaluator
    - iv. Document (e.g., lesson plans, data, artifacts) submission with feedback
    - v. Document (e.g., lesson plans, data, artifacts) submission without feedback
    - vi. Observations of a colleague
    - vii. Formal meetings with colleagues (e.g., data team meetings, professional learning community)
    - viii. Formal coaching or mentoring by non-evaluator
    - ix. A school- or district-based workshop during the academic year
    - x. A summer workshop
    - xi. A face-to-face course located outside of your district
    - xii. An online course or webinar
    - xiii. Attended a conference
  - b. What was the duration?
    - i. One day or less
    - ii. Two or more consecutive days
    - iii. Two or more days spread over several weeks
    - iv. Two or more days spread over several months (e.g., a semester)
    - v. Ongoing (no definitive end date)
  
6. Which of the following indicators did the professional development address? (select all that apply)
  - a. Indicator A — Instruction
  - b. Indicator B — Learning Environment
  - c. Indicator C — Cultural Proficiency
  - d. Indicator D — Expectations
  - e. I don't know
  
7. Did you engage in any professional development or support related to Standard 3: Family and Community Engagement?
  - a. Yes
  - b. No

8. For each type of professional development or support listed related to Standard 3: Family and Community Engagement, indicate whether you engaged in that type of professional development and indicate its duration.
  - a. Did you engage in this type of professional development? (Yes/No)
    - i. Tried specific strategies recommended by evaluator
    - ii. Read resources recommended by evaluator
    - iii. Regular meetings with evaluator
    - iv. Document (e.g., lesson plans, data, artifacts) submission with feedback
    - v. Document (e.g., lesson plans, data, artifacts) submission without feedback
    - vi. Observations of a colleague
    - vii. Formal meetings with colleagues (e.g., data team meetings, professional learning community)
    - viii. Formal coaching or mentoring by non-evaluator
    - ix. A school- or district-based workshop during the academic year
    - x. A summer workshop
    - xi. A face-to-face course located outside of your district
    - xii. An online course or webinar
    - xiii. Attended a conference
  - b. What was the duration?
    - i. One day or less
    - ii. Two or more consecutive days
    - iii. Two or more days spread over several weeks
    - iv. Two or more days spread over several months (e.g., a semester)
    - v. Ongoing (no definitive end date)
  
9. Which of the following indicators did the professional development address? (select all that apply)
  - a. Indicator A — Engagement
  - b. Indicator B — Collaboration
  - c. Indicator C — Communication
  - d. I don't know
  
10. Did you engage in any professional development or support related to Standard 4: Professional Culture?
  - a. Yes
  - b. No
  
11. For each type of professional development or support listed related to Standard 4: Professional Culture, indicate whether you engaged in that type of professional development and indicate its duration.
  - a. Did you engage in this type of professional development? (Yes/No)
    - i. Tried specific strategies recommended by evaluator
    - ii. Read resources recommended by evaluator
    - iii. Regular meetings with evaluator
    - iv. Document (e.g., lesson plans, data, artifacts) submission with feedback
    - v. Document (e.g., lesson plans, data, artifacts) submission without feedback
    - vi. Observations of a colleague
    - vii. Formal meetings with colleagues (e.g., data team meetings, professional learning community)

- viii. Formal coaching or mentoring by non-evaluator
  - ix. A school- or district-based workshop during the academic year
  - x. A summer workshop
  - xi. A face-to-face course located outside of your district
  - xii. An online course or webinar
  - xiii. Attended a conference
- b. What was the duration?
- i. One day or less
  - ii. Two or more consecutive days
  - iii. Two or more days spread over several weeks
  - iv. Two or more days spread over several months (e.g., a semester)
  - v. Ongoing (no definitive end date)
12. Which of the following indicators did the professional development address? (select all that apply)
- a. Indicator A — Reflection
  - b. Indicator B — Professional Growth
  - c. Indicator C — Collaboration
  - d. Indicator D — Decision-Making
  - e. Indicator E — Shared Responsibility
  - f. Indicator F — Professional Responsibilities
  - g. I don't know
13. How strongly do you agree or disagree with each of the following statements about the range of professional development or support you engaged in between May 2013 and May 2014? (Strongly agree, agree, disagree, strongly disagree)
- a. The professional development or support was relevant to my learning needs.
  - b. The professional development or support was relevant to the learning needs identified by my evaluator.
  - c. The professional development or support has helped me be a better teacher.
  - d. The professional development or support has helped me to better support my students in the classroom.
14. Were all of the professional development or support needs identified with your evaluator met during the May 2013–May 2014 timeframe?
- a. Yes
  - b. No (please explain what professional development or support needs you had that were not met)

## **Appendix G. Interview protocols**

Appendix G includes the teacher and evaluator interview protocols that were used with 10 interviewees for context information.

### **Teacher interview**

1. Can you describe in your own words what the needs were that were identified in the prescription/prescriptions you received in the last year?
2. How did you and your evaluator approach the task of writing the prescription?
3. What were you supposed to do to address the needs identified in the prescription? Was it clear what you needed to do to fulfill the prescription and address the needs identified in the prescription/s?
4. What have you done to fulfill the expectations of the prescription?
5. How well aligned was the professional development you received to the prescription? (Please provide an example of the extent of alignment or lack of alignment between the prescription and the PD.)
6. How did your experience with the prescription process compare to that of your colleagues?
7. Did you select any additional professional development that was not assigned in the prescription but that was relevant to your identified needs?
8. What has been most beneficial in helping to address the need that was identified in the prescription?
9. What types of professional development would have been helpful, if any, that were not available to you?
10. To what do you attribute any lack of available and relevant professional development?

### **Principal interview**

1. Did you write any prescription/s for teachers in your school in the 2013/14 school year? How many? What was the nature of the prescriptions you created?
2. How did you approach the task of writing the prescriptions? How did you know what you were supposed to do with regard to developing the prescriptions (guidance, direction, etc.)? (How did you decide what standard/indicator to label the prescription?)
3. What do you understand your role and responsibilities to be related to the writing and monitoring of the prescriptions?

4. Consider one or more of your teachers on prescriptions: how did you select the professional development or other activities for the teachers as part of the prescription?
5. What has the teacher/these teachers done to fulfill the expectations of their prescriptions?
6. How well aligned was the professional development the teachers received to the prescriptions? (Please provide an example of the extent of alignment or lack of alignment between the prescriptions and the PD.)
7. Did the teacher/these teachers satisfactorily address the need identified in the prescription/s?
8. If yes to #7, what has been most beneficial in helping to address the need/s that was/were identified in the prescription?
9. If no to #7, to what do you attribute the lack of improvement?
10. What types of professional development would have been helpful, if any, that were not available to the teacher/s?
11. To what do you attribute any lack of available and relevant professional development?

## References

- Behrstock-Sherratt, E., & Jacques, C. (2012). *Aligning evaluation results and professional development: Driving systemic human capital management reform*. Washington, DC: U.S. Department of Education Teacher Incentive Fund. Retrieved September 12, 2013, from [http://tpep-wa.org/wp-content/uploads/35492\\_TIF\\_HCM\\_Reform\\_FINALpercent5B1percent5D.pdf](http://tpep-wa.org/wp-content/uploads/35492_TIF_HCM_Reform_FINALpercent5B1percent5D.pdf).
- Blank, R. K., & de las Alas, N. (2009). *Effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders*. Washington, DC: Council of Chief State School Officers.
- Chambers, J. G., Lam, I., & Mahitivanichcha, K. (2008). *Examining context and challenges in measuring investment in professional development: A case study of six school districts in the Southwest Region* (Issues & Answers Report, REL 2008–037). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. <http://ies.ed.gov/ncee/edlabs>
- Coggsall, J. (2012). *Toward the effective teaching of new college- and career-ready standards: Making professional learning systematic*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved September 5, 2014, from <http://www.gtlcenter.org/sites/default/files/docs/TowardEffectiveTeaching.pdf>.
- Coggsall, J., Rasmussen, C., Colton, A., Milton, J., & Jacques, C. (2012). *Generating teacher effectiveness: The role of job-embedded professional learning in teacher evaluation*. Denver, CO: National Comprehensive Center for Teacher Quality. Retrieved September 12, 2013, from <http://www.gtlcenter.org/sites/default/files/docs/GeneratingTeachingEffectiveness.pdf>.
- Croft, A., Coggsall, J., Dolan, M., Powers, E., & Killion, J. (2010). *Job-embedded professional development: What it is, who is responsible, and how to get it done well* (Issue Brief). Washington, DC, and Oxford, OH: National Comprehensive Center for Teacher Quality, Mid-Atlantic Comprehensive Center, and National Staff Development Council. Retrieved September 5, 2014, from <http://www.gtlcenter.org/sites/default/files/docs/JEPDpercent20Issuepercent20Brief.pdf>.
- Curtis, R. (2011). *District of Columbia public schools: Defining instructional expectations and aligning accountability and support*. Washington, DC: The Aspen Institute. Retrieved July 18, 2012, from [http://www.aspeninstitute.org/sites/default/files/content/docs/education/AI\\_DCPS\\_teacherpercent20evaluation.pdf](http://www.aspeninstitute.org/sites/default/files/content/docs/education/AI_DCPS_teacherpercent20evaluation.pdf).
- Curtis, R., & Wiener, R. (2012). *Means to an end: A guide to developing teacher evaluation systems that support growth and development*. Washington, DC: The Aspen Institute. Retrieved September 5, 2014, from <http://www.aspeninstitute.org/publications/means-end-guide-developing-teacher-evaluation-systems-support-growth-development>.

- Daley, G., & Kim, L. (2010). *A teacher evaluation system that works: A working paper*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved June 27, 2011, from [http://www.tapsystem.org/publications/wp\\_eval.pdf](http://www.tapsystem.org/publications/wp_eval.pdf).
- Danielson, C. (2012). It's your evaluation—Collaborating to improve teacher practice. *Education Digest: Essential Readings Condensed for Quick Review*, 77(8), 22–27. <http://eric.ed.gov/?id=EJ982506>
- Darling-Hammond, L., & Stanford Center for Opportunity Policy in Education. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Palo Alto, CA: Stanford Center for Opportunity Policy in Education. Retrieved September 5, 2014, from [https://edpolicy.stanford.edu/sites/default/files/publications/creating-comprehensive-system-evaluating-and-supporting-effective-teaching\\_1.pdf](https://edpolicy.stanford.edu/sites/default/files/publications/creating-comprehensive-system-evaluating-and-supporting-effective-teaching_1.pdf).
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Van Petegem, P. (2013). How may teacher evaluations have an impact on professional development? A multi-level analysis. *Teaching & Teacher Education*, 36(1), 1–11.
- Desimone, L., Porter, A., Garet, M., Yoon, K., & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, 24(2), 81–112. Retrieved September 12, 2013, from <http://www.jstor.org/stable/3594138>.
- Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915–945. Retrieved September 12, 2013, from <http://www.jstor.org/stable/3202507>.
- Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., & Whitehurst, G. (2011). *Passing muster: Evaluating teacher evaluation systems*. Washington, DC: Brookings Institution.
- Goe, L., Biggers, K., & Croft, A. (2012). *Linking teacher evaluation to professional development: Focusing on improving teaching and learning*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved July 1, 2013, from <http://gtl.airprojects.org/sites/default/files/docs/LinkingTeacherEval.pdf>.
- Goe, L., Holdheide, L., & Miller, T. (2014). *Practical guide to designing comprehensive teacher evaluation systems*. (Revised edition.) Washington, DC: Center on Great Teachers & Leaders.
- Gordon, R., Kane, T. J., & Staiger, D. O. (2006). *Identifying effective teachers using performance on the job*. Washington, DC: Brookings Institution. Retrieved November 17, 2010, from [http://www.brookings.edu/~media/Files/rc/papers/2006/04education\\_gordon/200604hamilton\\_1.pdf](http://www.brookings.edu/~media/Files/rc/papers/2006/04education_gordon/200604hamilton_1.pdf).
- Great Schools Partnership. (2014). Professional development. In S. Abbott (Ed.), *The glossary of education reform*. Retrieved July 16, 2015, from <http://edglossary.org/professional-development/>.

- Guskey, T., & Yoon, K. (2009). What works in professional development? *The Phi Delta Kappan*, 90(7), 495–500. Retrieved September 12, 2013, from <http://www.jstor.org/stable/20446159>.
- Hamilton, L., Steiner, E., Holtzman, D., Fulbeck, E., Robyn, A., & Poirier, J., et al. (2014). *Using teacher evaluation data to inform professional development in the intensive partnership sites*. RAND Working Paper. Retrieved May 19, 2015, from [http://www.rand.org/content/dam/rand/pubs/working\\_papers/WR1000/WR1033/RAND\\_WR1033.pdf](http://www.rand.org/content/dam/rand/pubs/working_papers/WR1000/WR1033/RAND_WR1033.pdf).
- Heneman, H. G. III, Milanowski, A., Kimball, S. M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay*. Policy Brief RB-45. Philadelphia, PA: Consortium for Policy Research in Education. Retrieved November 17, 2010, from <http://cpre.wceruw.org/publications/rb45.pdf>.
- Ho, A. D., & Kane, T. J. (2013). *The reliability of classroom observations by school personnel*. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved July 16, 2015, from [http://www.metproject.org/downloads/MET\\_Reliability%20of%20Classroom%20Observations\\_Research%20Paper.pdf](http://www.metproject.org/downloads/MET_Reliability%20of%20Classroom%20Observations_Research%20Paper.pdf).
- Hull, J. (2013). *Trends in teacher evaluation: How states are measuring teacher performance*. Alexandria, VA: Center for Public Education. Retrieved September 5, 2014, from <http://www.centerforpubliceducation.org/gm-node/2408.aspx>.
- Kersting, N. B., Chen, M., & Stigler, J. W. (2013). Value-added teacher estimates as part of teacher evaluations: Exploring the effects of data and model specifications on the stability of teacher value-added scores. *Education Policy Analysis Archives*, 21(7), 1–39. Retrieved July 3, 2013, from <http://epaa.asu.edu/ojs/article/view/1167>.
- Learning Forward. (2014). *Definition of professional development*. Retrieved July 16, 2015, from <http://learningforward.org/who-we-are/professional-learning-definition#DefinitionResources>.
- Learning Point Associates. (2010). *Emerging trends reflected in the state phase 1: Race to the Top applications*. Evaluating Teacher Effectiveness. Naperville, IL: Learning Point Associates. <http://www.eric.ed.gov/PDFS/ED509942.pdf>
- Measures of Effective Teaching Project. (2010). *Learning about teaching: Initial findings from the measures of effective teaching project*. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved July 5, 2011, from <http://s3.documentcloud.org/documents/18327/met-research-paper.pdf>.
- Measures of Effective Teaching Project. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved January 25, 2012, from [http://metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).
- Milanowski, A. T., & Heneman, H. (2001). Assessment of teacher reactions to a standards-based teacher evaluation system: A pilot study. *Journal of Personnel Evaluation in Education*, 15(3), 193–212.

- National Council on Teacher Quality. (2013). *State of the states 2013: Teacher effectiveness policies*. Washington, DC: National Council on Teacher Quality. Retrieved September 30, 2014, from [http://www.nctq.org/dmsView/State\\_of\\_the\\_States\\_2013\\_Using\\_Teacher\\_Evaluations\\_NCTQ\\_Report](http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report).
- Odden, A. (2011). The dollars and sense of comprehensive professional learning. *Journal of Staff Development*, 32(4), 26–32. Retrieved September 12, 2013, from [http://tpep-wa.org/wp-content/uploads/35492\\_TIF\\_HCM\\_Reform\\_FINALpercent5B1percent5D.pdf](http://tpep-wa.org/wp-content/uploads/35492_TIF_HCM_Reform_FINALpercent5B1percent5D.pdf).
- Parise, L. M., & Spillane, J. P. (2010). Teacher learning and instructional change: How formal and on-the-job learning predict change in elementary school teachers' practice. *The Elementary School Journal*, 110(3), 323–346.
- Penuel, W. R., Fishman, B., Yamaguchi, R., & Gallagher, L. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44(4), 921–958.
- Sartain, L., Stoelinga, S. R., & Brown, E. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Chicago, IL: Consortium on Chicago School Research, University of Chicago Urban Education Institute. Retrieved January 23, 2012, from <https://consortium.uchicago.edu/sites/default/files/publications/Teacher%20Eval%20Report%20FINAL.pdf>.
- Supovitz, J. A., Mayer, D., & Kahle, J. B. (2000). The longitudinal impact of inquiry-based professional development on teaching practice. *Educational Policy*, 14(3), 331–356.
- Taylor, E. S., & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (NBER Working Paper No. 16877). Cambridge, MA: National Bureau of Economic Research. <http://eric.ed.gov/?id=ED517378>
- Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. (Education Sector Reports.) Washington, D.C.: Education Sector. Retrieved March 10, 2011, from [http://www.educationsector.org/sites/default/files/publications/RushToJudgment\\_ES\\_Jan08.pdf](http://www.educationsector.org/sites/default/files/publications/RushToJudgment_ES_Jan08.pdf).
- Tuytens, M., & Devos, G. (2011). Stimulating professional learning through teacher evaluation: An impossible task for the school leader? *Teaching & Teacher Education*, 27(5), 891–899. <http://eric.ed.gov/?id=EJ924136>
- Tuytens, M., & Devos, G. (2014). How to activate teachers through teacher evaluation? *School Effectiveness & School Improvement*, 25(4), 509–530. <http://eric.ed.gov/?id=EJ1034470>
- U.S. Department of Education. (2011). *Guidance on fiscal year 2010 School Improvement Grants*. Washington, DC: Office of Elementary and Secondary Education. Retrieved June 25, 2012, from <http://www2.ed.gov/programs/sif/sigguidance02232011.pdf>.

- U.S. Department of Education. (2013). *ESEA flexibility*. Retrieved July 3, 2013, from <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>.
- Wallace, M. R. (2009). Making sense of the links: Professional development, teacher practices, and student achievement. *Teachers College Record*, 111(2), 573–596.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: The New Teacher Project. Retrieved July 5, 2011, from <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>.
- Weiss, I., & Pasley, J. (2006). *Scaling up instructional improvement through teacher professional development: Insights from the local systemic change initiative*. (CPRE Policy Brief No. RB-44.) Philadelphia, PA: Consortium for Policy Research in Education. <http://eric.ed.gov/?id=ED493115>
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. <http://eric.ed.gov/?id=ED498548>

## The Regional Educational Laboratory Program produces 7 types of reports



### **Making Connections**

Studies of correlational relationships



### **Making an Impact**

Studies of cause and effect



### **What's Happening**

Descriptions of policies, programs, implementation status, or data trends



### **What's Known**

Summaries of previous research



### **Stated Briefly**

Summaries of research findings for specific audiences



### **Applied Research Methods**

Research methods for educational settings



### **Tools**

Help for planning, gathering, analyzing, or reporting data or research