Making Connections

November 2016

Teachers' responses to feedback from evaluators: What feedback characteristics matter?

Trudy L. Cherasaro
R. Marc Brodersen
Marianne L. Reale
David C. Yanoski
Marzano Research

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL CENTRAL
Regional Educational Laboratory
At Marzano Research

U.S. Department of Education
John B. King, Jr., *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*
Sandra Garcia, *Project Officer*

REL 2017–190

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.

November 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0007 by Regional Educational Laboratory Central administered by Marzano Research. 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:

Cherasaro, T. L., Brodersen, R. M., Reale, M. L., & Yanoski, D. C. (2016). *Teachers' responses to feedback from evaluators: What feedback characteristics matter?* (REL 2017–190). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. 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

In response to initiatives to increase educator effectiveness as directed through flexibility waivers under the Elementary and Secondary Education Act, many states are implementing new teacher evaluation systems. Those states are also seeking information about how evaluators can best use evaluation findings to provide individualized feedback to teachers to improve both teaching and learning. Data from Regional Educational Laboratory Central's Examining Evaluator Feedback Survey were used to analyze teachers' perceptions of feedback provided as part of their district's teacher evaluation system as well as their ratings of the importance of various characteristics of feedback in their response to feedback. The study team then explored how characteristics of feedback and response to feedback are interrelated.

Correlational analysis finds that teachers' responses to feedback are related to their perceptions of four characteristics: the usefulness of the feedback, the accuracy of the feedback, the credibility of their evaluator, and their access to resources. Structural equation modeling analysis suggests that in responding to feedback, teachers' perceptions of the usefulness of the feedback and the credibility of their evaluator could be more important than their perceptions of the accuracy of the feedback and their access to resources.

Results from this study may be helpful in prioritizing evaluation needs at both the state and district levels for training and guidance on providing feedback. They may also help inform states of additional data needed to improve understanding of how feedback is used and what impact it can have on teacher performance.

Contents

Tables

Why this study?

The importance of teacher effectiveness is well supported by studies that document variation in teachers' abilities to contribute to student achievement gains. All else being equal, students taught by some teachers experience greater achievement gains than do students taught by other teachers (Aaronson, Barrow, & Sander, 2007; Chetty, Friedman, & Rockoff, 2014; Konstantopoulos & Chung, 2010; Nye, Konstantopoulos, & Hedges, 2004; Wright, Horn, & Sanders, 1997). Increasing confirmation of the importance of teachers' contributions to student learning has led to an interest in identifying and supporting effective teachers through teacher evaluation systems, as evidenced by the focus on developing teacher evaluation systems in the Elementary and Secondary Education Act of 1965 (ESEA) flexibility waivers.[1] In a well designed, well implemented evaluation system, teachers receive sufficient evidence-based feedback to guide their reflection on their strengths and challenges in order to improve their practice (Coggshall, Rasmussen, Colton, Milton, & Jacques, 2012).

The purpose of this study is to support efforts to increase teacher effectiveness by examining how teachers value and use different aspects of the feedback they receive

One of the four principles of the ESEA flexibility waivers is supporting effective instruction and leadership. This principle requires states to develop and implement teacher evaluation and support systems that "provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development" (U.S. Department of Education, 2012, p. 6). By 2015 all but one state had requested an ESEA flexibility waiver (Center on Education Policy, n.d.).[2] All states in the Regional Educational Laboratory (REL) Central Region are developing or implementing new teacher evaluation systems that focus on teacher development. As of 2015 four states had an approved ESEA flexibility waiver (Colorado, Kansas, Missouri, and South Dakota), two states had submitted waivers that are pending approval (Nebraska and Wyoming), and one state withdrew its waiver (North Dakota). Even states whose waiver applications were withdrawn, revoked, or denied still have plans to implement new evaluation systems with a focus on teacher development.

This study was conducted before passage of the Every Student Succeeds Act, which has no requirements related to teacher evaluation systems. Because many states have already begun implementing new systems and therefore have changed their policies, it is likely that these states will continue in these efforts. Additionally, whether within or outside of formal evaluation, providing teachers with feedback based on their performance could help improve education systems. So the study's findings may still provide useful information.

As states and districts develop and implement teacher evaluation systems, they are exploring ways to use evaluation findings to provide individualized feedback that will facilitate improved teaching and learning practices, which will lead to better student performance (Kane & Staiger, 2012). As these systems have been developed, state and district administrators in the REL Central Region have expressed a growing interest in learning more about the quality and usefulness of the feedback teachers receive. Administrators seek this information to inform future efforts to provide teachers useful feedback that will help guide their professional development.

The purpose of this study is to support efforts to increase teacher effectiveness by examining how teachers value and use different aspects of the feedback they receive. The findings are designed to help stakeholders improve the provision and use of feedback in teacher evaluation systems by better understanding teachers' perceptions of feedback and how they

respond to it. This information may help states and districts prioritize training and guidance needs in providing feedback.

What the study examined

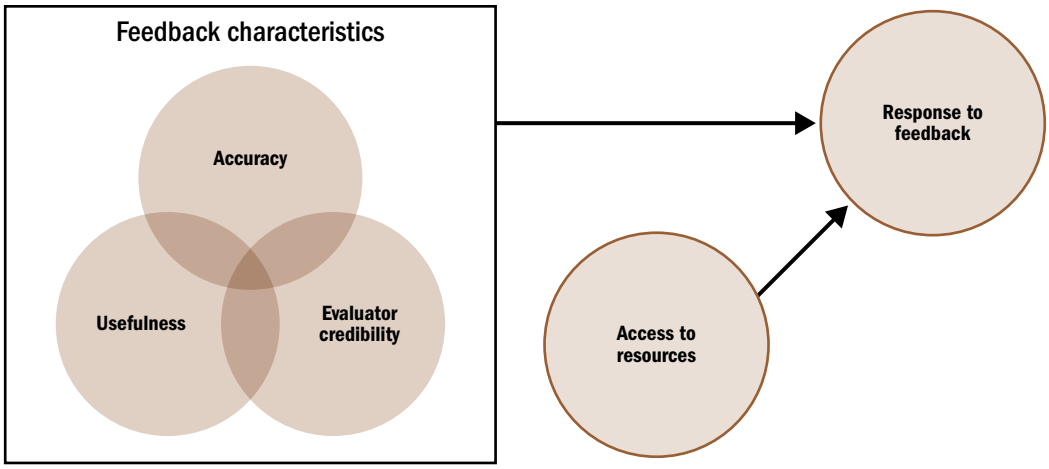
Although many characteristics may influence the use of feedback from evaluators, the current study used a theoretical model that includes several key characteristics that were identified in the literature (figure 1). The study sought to understand how response to feedback from evaluators is influenced by teachers' perceptions of four interrelated characteristics:

- *Usefulness of the feedback.* The perception that the feedback provides specific recommendations, is provided in a timely manner, and is provided frequently (Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008; Goldhaber & Brewer, 2000; Hemmeter, Snyder, Kinder, & Artman, 2011; Monk & King, 1994; Johnson, 2000; Scheeler, Ruhl, & McAfee, 2004; Supovitz, 2012; Tuytens & Devos, 2011).
- *Accuracy of the feedback.* The perception that the feedback accurately represents performance (Kinicki, Prussia, Wu, & McKee-Ryan, 2004).
- *Evaluator credibility.* The perception that the evaluator has the knowledge and understanding to give valuable feedback (Coggshall et al., 2012; Kinicki et al., 2004; Tuytens & Devos, 2011).
- *Access to resources.* The extent to which the teacher believes that he or she will have access to expert teachers, mentors, coaches, or other professional development opportunities needed to develop the skills or knowledge to respond to the feedback (Desimone, Porter, Garet, Yoon, & Birman, 2002; Garet, Porter, Desimone, Birman, & Yoon, 2001; Parise & Spillane, 2010).

The study sought to understand how response to feedback from evaluators is influenced by teachers' perceptions of four interrelated characteristics: usefulness of the feedback, accuracy of the feedback, evaluator credibility, and access to resources

Any changes in teacher practices or any support that teachers sought in order to make such changes was considered response to feedback. This theoretical model was developed based on a review of research on performance feedback inside and outside of the field of education, including Ilgen, Fisher, and Taylor's (1979) seminal work on effective feedback.

Figure 1. Theoretical model for performance feedback in teacher evaluation systems



Source: Authors' construction.

This study addresses three research questions:

1. What are teachers' perceptions of the usefulness and accuracy of the feedback they received as a part of their evaluation system and what are their perceptions of their evaluator's credibility and their access to resources related to the feedback?

2. How are the perceptions of the usefulness of feedback, the accuracy of feedback, evaluator credibility, and access to resources interrelated?

3. How are the usefulness of feedback, the accuracy of feedback, evaluator credibility, and access to resources related to the response to feedback?

A teacher survey was used to gather data to address the research questions. The survey was administered online in spring 2015 and asked teachers to reflect on the feedback they had received from their evaluator throughout the 2014/15 school year. The analyses to address the questions were descriptive (research question 1) and correlational (research questions 2 and 3), so causal inferences cannot be drawn about how teachers' perceptions of feedback or how access to resources affect their response to feedback. The data and methods used in this study are summarized in box 1, with additional detail provided in appendix A.

A teacher survey, administered online in spring 2015, asked teachers to reflect on the feedback they had received from their evaluator throughout the 2014/15 school year

Box 1. Data and methods

Data

This study included a purposive sample of teachers from seven school districts in two states in the Regional Educational Laboratory Central Region. The districts were involved in the pilot test of their state's new teacher evaluation system (which focused on providing individualized feedback) for one to three years before data collection. The sample comprised 317 preK–12 teachers who were being evaluated using the district's new teacher evaluation system; who had direct student contact in a classroom setting; who were from urban locales, rural locales, and small towns; and who taught various subject areas, including English language arts, math, science, social studies, and noncore subjects such as physical education, art, and technology. The sampled teachers also taught a range of student populations, including English learner and special education students (see appendix A).

The findings are based on responses to the Examining Evaluator Feedback Survey (Cherasaro et al., 2015), which was administered online during spring 2015 and asked teachers to reflect on the feedback they had received throughout the 2014/15 school year from their evaluator and to answer questions about who provided the feedback; how often it was provided; their perceptions of the usefulness and accuracy of the feedback, evaluator credibility, access to resources, and response to the feedback; the importance of feedback characteristics and their beliefs about instructional improvement; and their demographic characteristics (see table B1 in appendix B). The survey had a response rate of 76.7 percent, with 243 of 317 teachers responding. Teachers indicated their level of agreement or disagreement with a series of statements on a five-point scale and indicated how important various aspects of feedback were to them. The survey was piloted in a previous study that generated evidence of reliability and validity (Cherasaro et al., 2015; see appendix B).

(continued)

Box 1. Data and methods *(continued)*

Methods

To examine teachers' perceptions of feedback (research question 1), frequencies and percentages were calculated for survey responses related to the feedback characteristics (survey questions 5–9) and the importance of feedback characteristics (survey questions 10–13). Data were analyzed for teachers who reported that they had received written feedback at least once or had at least one feedback conversation (95 percent of the 243 survey respondents). To examine the relationships among characteristics of feedback (research question 2) and how characteristics of feedback are related to response to feedback (research question 3), the study team calculated correlations between various pairs of variables.

First, a scale score was created for each domain (usefulness, accuracy, evaluator credibility, access to resources, and response to feedback) by averaging responses for the set of questions within the domain. Bivariate correlations between the scale scores for each domain were examined. The correlational analysis was followed by structural equation modeling techniques to examine the relationships between the variables after accounting for other relationships. Structural equation modeling was used to examine how teachers' perceptions of the usefulness of feedback, accuracy of feedback, and evaluator credibility related to one another, as well as how these variables and teachers' access to resources influenced the actions they took in response to the feedback they received. Structural equation modeling was used to test the theoretical model and make adjustments to identify a model that was the most parsimonious and best fit the data (see appendix B).

What the study found

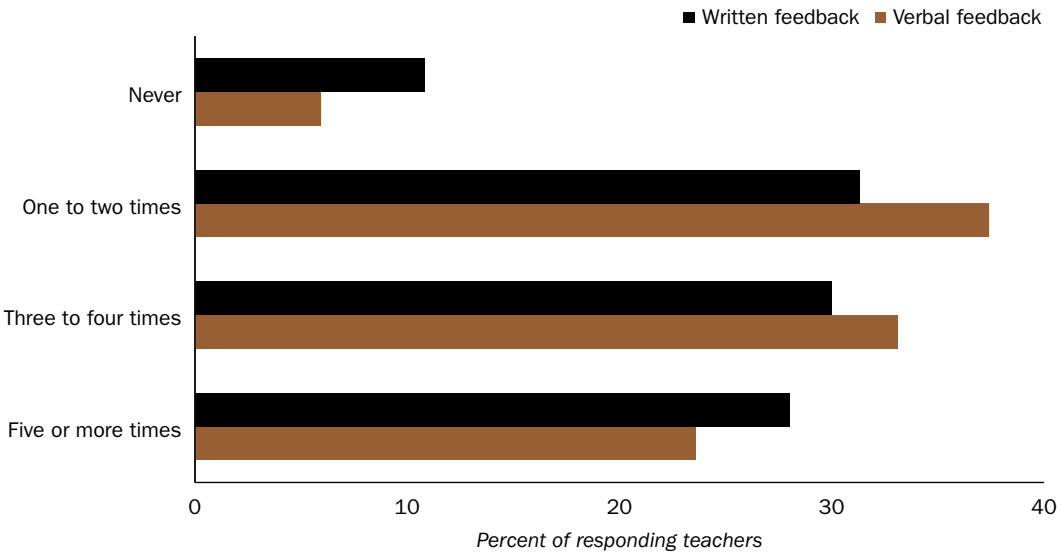
Teachers who responded to the survey were most often evaluated by their principal (70 percent) or assistant principal (27 percent), and most received written and verbal feedback between one and four times per year

This section first discusses the findings related to research question 1, how teachers perceive their feedback, by presenting teachers' level of agreement with and rating of importance of statements related to feedback usefulness, accuracy, evaluator credibility, access to resources, and response to feedback. It then discusses the findings related to research questions 2 and 3, which describes the relationship among feedback usefulness, accuracy, evaluator credibility, access to resources, and response to feedback. To understand the type of feedback teachers are reporting on, it is important to note that teachers who responded to the survey were most often evaluated by their principal (70 percent) or assistant principal (27 percent) and that most of these teachers received written and verbal feedback between one and four times per year (figure 2).

Most teachers agreed that the feedback they received was useful and accurate and that their evaluator was credible

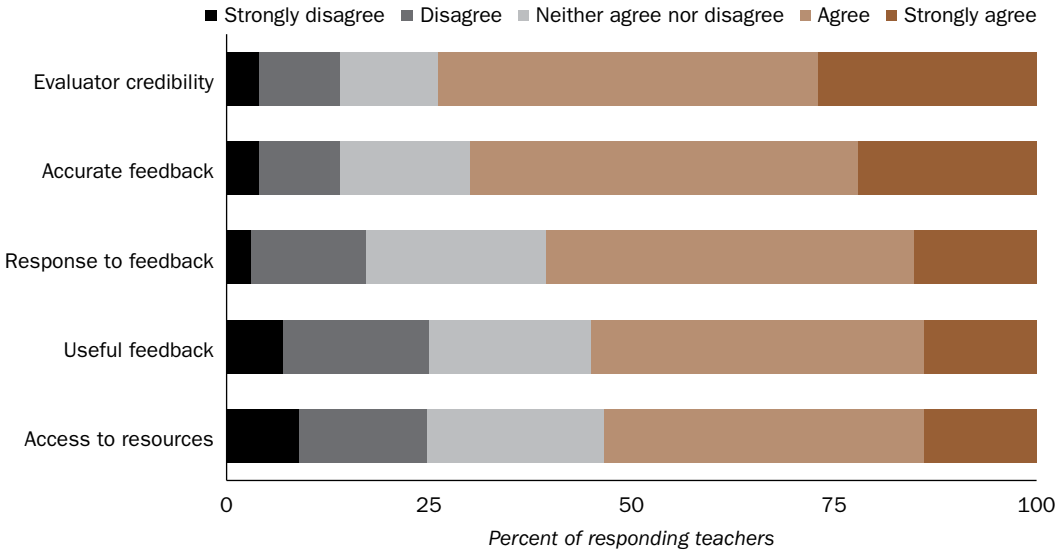
Some 74 percent of teachers agreed or strongly agreed that their evaluator was credible, 70 percent agreed or strongly agreed that the feedback they received was accurate, and 55 percent of teachers agreed or strongly agreed that the feedback they received was useful (figure 3). More than half of teachers also agreed or strongly agreed that they responded to the feedback they received (60 percent) and that they had access to resources related to the feedback they received (54 percent).

Figure 2. Most teachers reported that they received written and verbal feedback between one and four times a year during 2014/15



Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Figure 3. Most teachers agreed that their evaluator was credible and that the feedback they received during 2014/15 was accurate



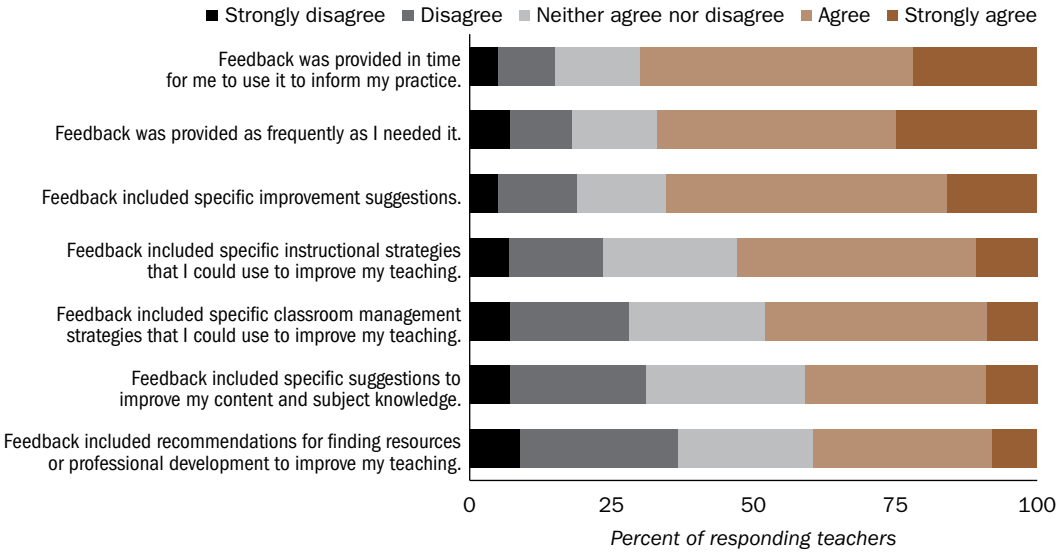
Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Some 74 percent of teachers agreed or strongly agreed that their evaluator was credible, 70 percent agreed or strongly agreed that the feedback they received was accurate, and 55 percent of teachers agreed or strongly agreed that the feedback they received was useful

More than 60 percent of teachers agreed or strongly agreed with three statements related to the usefulness of the feedback they received: that feedback was timely (70 percent), was frequent (67 percent), and included specific suggestions for improvement (66 percent; figure 4).

Seventy-five percent or more of teachers agreed or strongly agreed with two statements related to the accuracy of the feedback they received: that the observations that informed

Figure 4. Most teachers agreed that the feedback they received during 2014/15 was timely, was frequent, and included specific suggestions for improvement



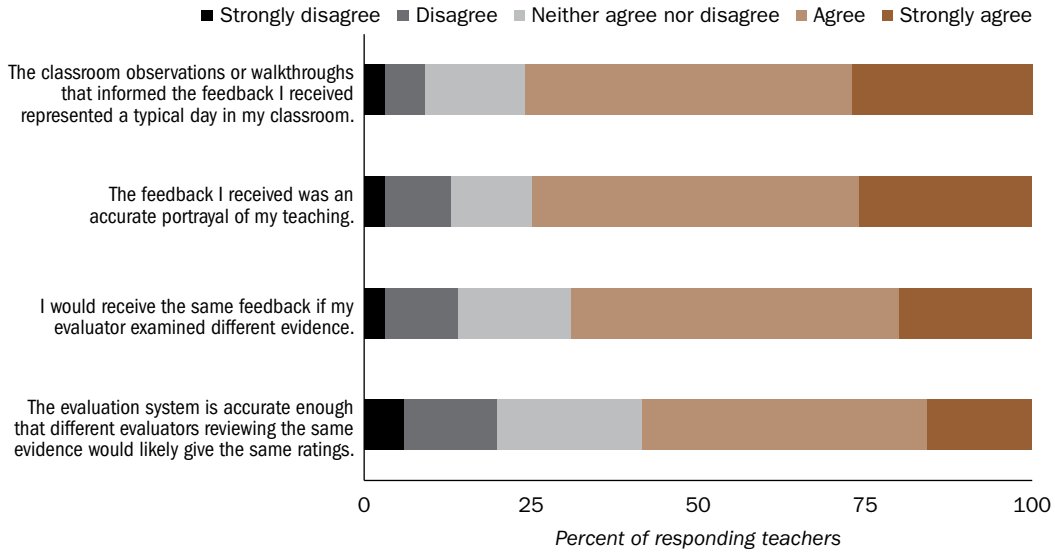
Some 70 percent of teachers agreed or strongly agreed that feedback was timely

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

the feedback represented a typical day in their classroom (76 percent) and that the feedback accurately portrayed their teaching (75 percent; figure 5).

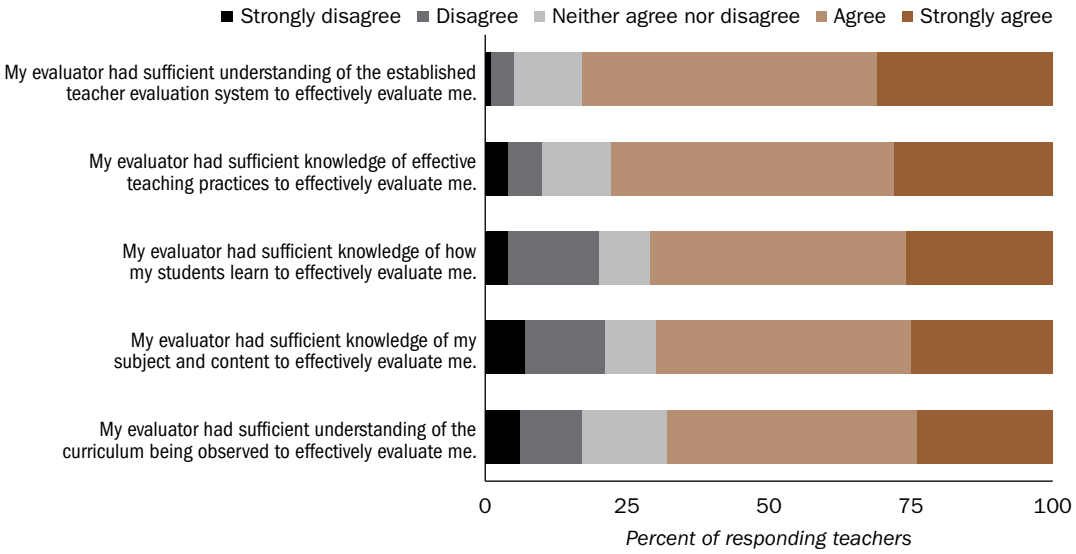
At least 68 percent of teachers agreed or strongly agreed with all statements related to the credibility of their evaluator (figure 6). The highest level of agreement (83 percent) was with the statement that their evaluator had sufficient understanding of the established teacher evaluation system to effectively evaluate them.

Figure 5. Most teachers agreed that the observations that informed the feedback they received during 2014/15 represented a typical day in their classroom



Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Figure 6. Most teachers agreed that their evaluator was credible, 2014/15

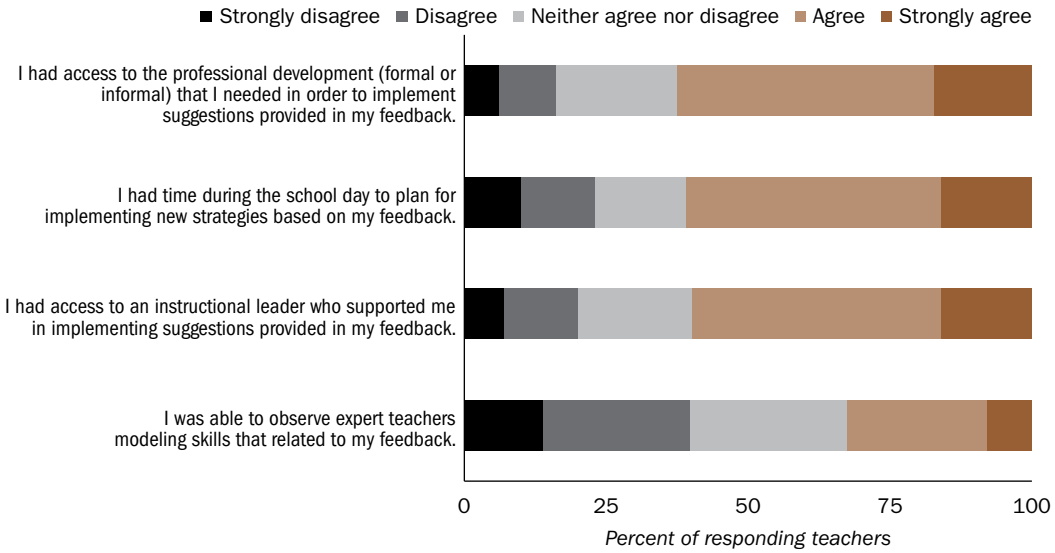


Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Some 62 percent of teachers agreed or strongly agreed that they had access to professional development

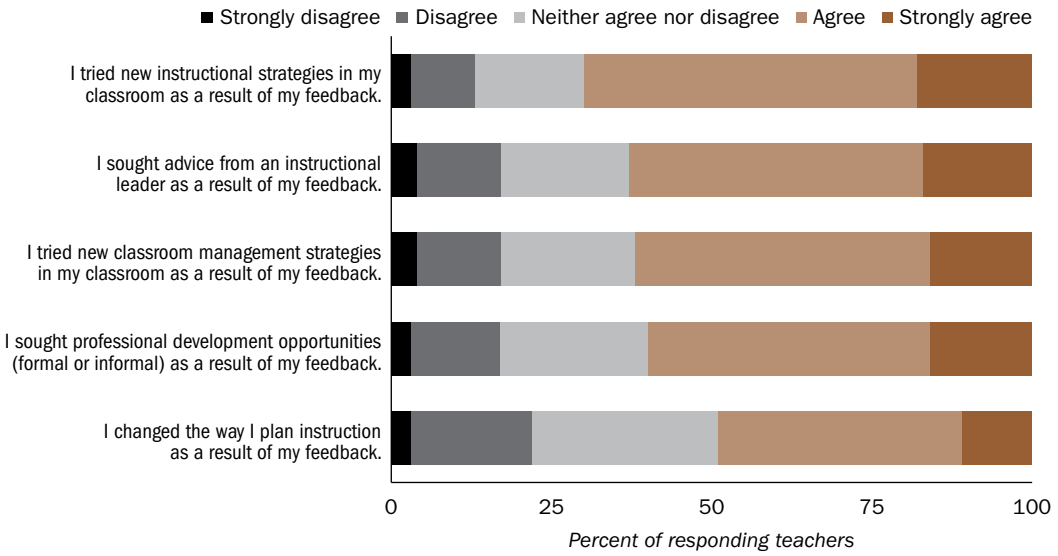
Sixty percent or more of teachers agreed or strongly agreed with three statements related to their access to resources: that they had access to professional development (62 percent), that they had time during the school day to plan for implementing new strategies (61 percent), and that they received support from an instructional leader (60 percent). But only 33 percent agreed or strongly agreed that they were able to observe expert teachers who modeled skills related to the feedback (figure 7).

Figure 7. Most teachers agreed that they had access to many resources, but few teachers agreed that they were able to observe expert teachers modeling skills related to the feedback, 2014/15



Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Figure 8. Most teachers agreed that they tried new instructional strategies as a result of feedback, 2014/15



Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

More than 60 percent of teachers indicated that feedback provided in an appropriate timeframe, feedback that was an accurate portrayal of their teaching, feedback that used observational data from a typical day in the classroom, and having time during the day to plan for implementing feedback were very important or critical in their response to feedback

Sixty percent or more of teachers responded to feedback in one of four ways: trying new instructional strategies (70 percent), seeking advice from an instructional leader (63 percent), trying new classroom management strategies (62 percent), and seeking professional development opportunities (60 percent; figure 8).

Teachers indicated that evaluator credibility was the most important characteristic affecting their response to feedback

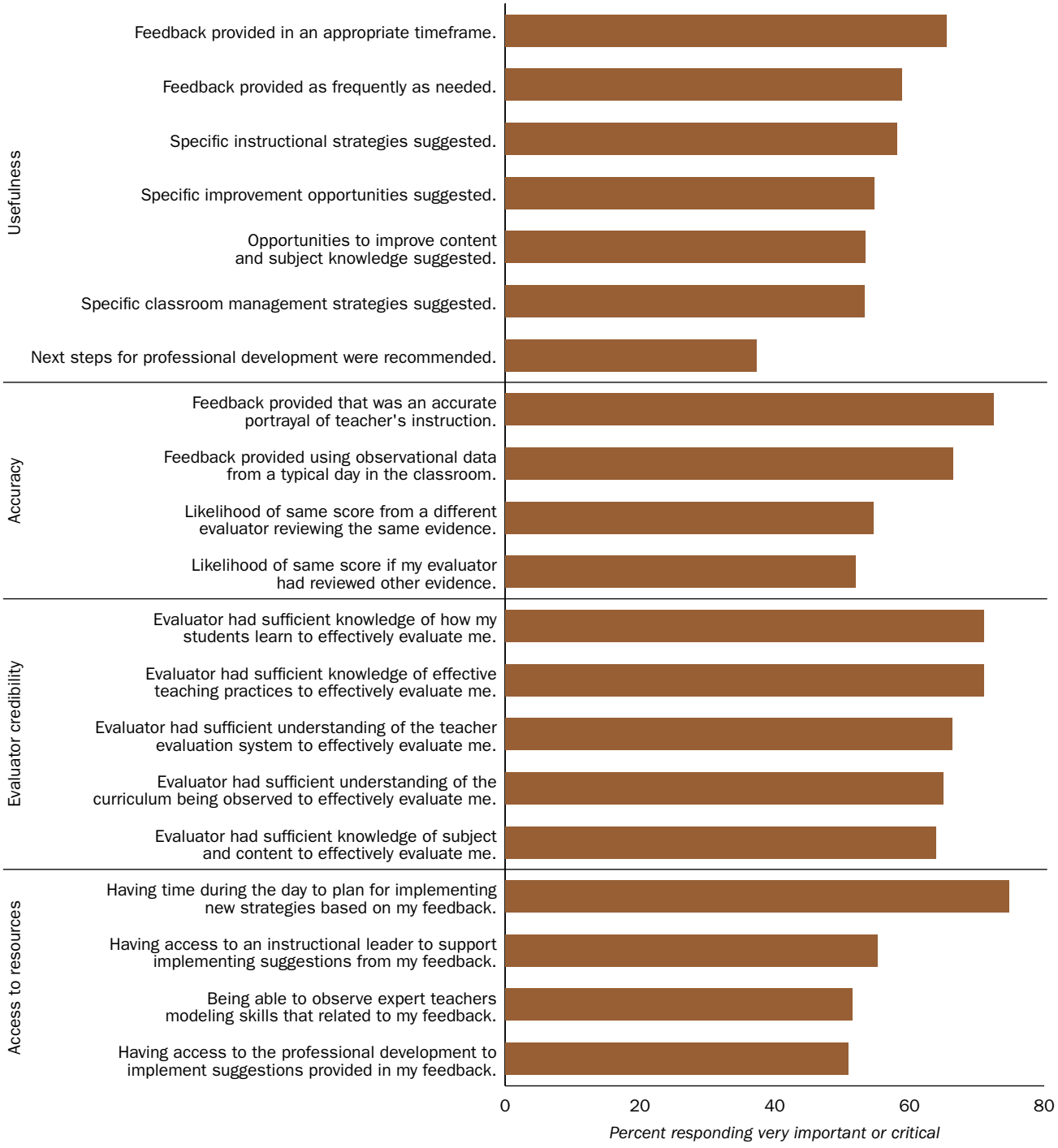
More than 60 percent of teachers rated every statement related to the credibility of their evaluator as very important or critical in their response to feedback (figure 9). Specific characteristics related to usefulness, accuracy, and access to resources were also perceived as important. More than 60 percent of teachers indicated that feedback provided in an appropriate timeframe, feedback that was an accurate portrayal of their teaching, feedback that used observational data from a typical day in the classroom, and having time during the day to plan for implementing feedback were very important or critical in their response to feedback.

Teachers' response to feedback is influenced by how useful they perceive it to be, which is influenced by how credible they perceive their evaluator to be

Bivariate correlational analysis found significantly positive correlations between all possible pairs of the following domains: usefulness of feedback, accuracy of feedback, evaluator credibility, access to resources, and response to feedback (table 1). How teachers respond to feedback is most closely related to how useful they perceive it to be.

Structural equation modeling analysis revealed a final model in which teachers' response to feedback was strongly correlated with the extent to which teachers perceived the feedback as useful, which in turn was strongly correlated with how credible they perceived their

Figure 9. Teachers indicated that they viewed evaluator credibility as the most important consideration in deciding how to respond to feedback, 2014/15



Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

evaluator to be, which was strongly correlated with how accurately they believed that the feedback represented their teaching (figure 10). The final model varied from the hypothesized model in that usefulness is determined partially by accuracy and evaluator credibility as opposed to usefulness, accuracy, and evaluator credibility being three independent but related aspects of perceived feedback quality (see appendix B). Although the correlational

Table 1. Teachers' response to feedback was related to the usefulness of the feedback, accuracy of the feedback, evaluator credibility, and access to resources, 2014/15

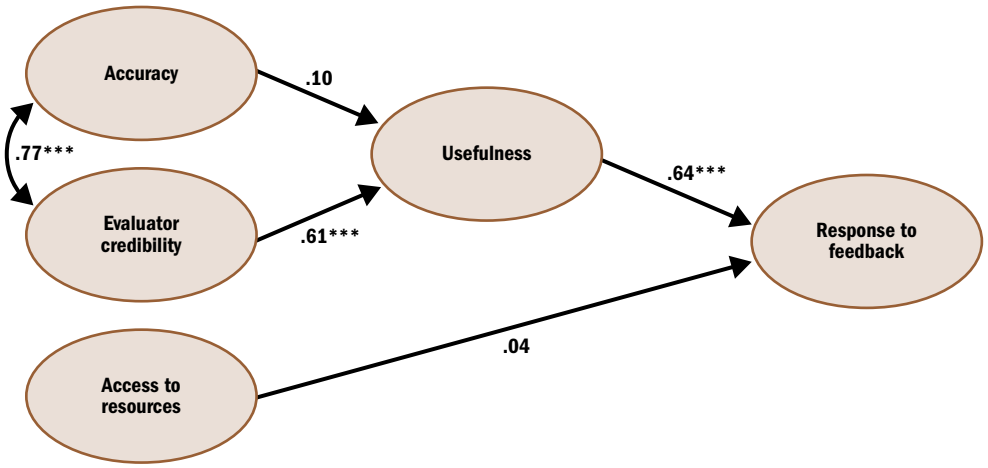
Domain	Accuracy of the feedback	Evaluator credibility	Access to resources	Response to feedback
Usefulness of the feedback	0.547**	0.672**	0.620**	0.596**
Accuracy of the feedback		0.708**	0.591**	0.182**
Evaluator credibility			0.617**	0.356**
Access to resources				0.403**

** Significant at $p < .01$.

Note: Table displays Pearson's *r* correlation coefficients for domain averages.

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Figure 10. Teachers' response to feedback was influenced by their perception of its usefulness, 2014/15



*** Significant at $p < .001$.

Note: Arrows represent a relationship between variables, with the direction of the arrow showing which variable influenced the other. The numbers above the arrows are STDYX path coefficients, which are regression coefficients that are standardized on both the independent and dependent variables. Each number represents the standardized amount of change in the dependent variable that would occur with one standardized standard deviation in the independent variable—or the strength, direction, and significance of the relationship. Larger numbers indicate stronger relationships, and positive numbers indicate that an increase in one variable is related to an increase in the other variable.

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

analysis revealed a significant positive correlation between accuracy and usefulness, the model suggests that after the relationship between evaluator credibility and usefulness is accounted for, accuracy has little additional influence on usefulness. In addition, although the correlational analysis revealed a strong positive correlation between access to resources and teachers' response to feedback, the relationship is not significant after the influence of the perceived usefulness of the feedback is accounted for.

Implications of the study findings

Teacher evaluation systems have the potential to support teachers in building their professional practice. The findings from this study provide important information about how teachers perceive and use feedback from evaluators. The results from the correlational analysis suggest that teachers' response to feedback is related to their perceptions of the usefulness of the feedback, the accuracy of the feedback, the credibility of their evaluator, and the resources they have access to. The results from the structural equation modeling analysis further suggest that teachers may be more likely to respond to feedback if they perceive that the feedback is useful than if they perceive that the feedback is not useful. And teachers may be more likely to perceive that the feedback is useful if they believe that their evaluator is credible than if they believe that their evaluator is not credible. The extent to which teachers believe that their evaluator is credible is strongly related to their perception of the accuracy of the feedback.

The results from the structural equation modeling analysis suggest that teachers may be more likely to respond to feedback if they perceive that the feedback is useful, and they may be more likely to perceive that the feedback is useful if they believe that their evaluator is credible

Although the findings do not support any conclusions about causality, they support a model of feedback on teacher performance that suggests characteristics to consider in training and supporting evaluators. Based on the findings, state and district education leaders might consider several actions:

- Reviewing evaluator training and guidance on feedback to teachers to identify ways to strengthen the usefulness of the feedback.
- Examining policies related to the usefulness of feedback or collecting data to identify potential barriers to providing useful feedback.
- Considering ways to ensure that feedback is frequent, is timely, and includes specific suggestions to improve content and subject knowledge, instructional strategies, classroom management strategies, and recommendations for finding resources or professional development opportunities.
- Targeting suggestions to improve content and subject knowledge and classroom management because more than half of teachers indicated that these suggestions were important for responding to feedback, but less than half said that these suggestions were provided in the feedback they received.
- Focusing on ways to build evaluator credibility because perceptions of evaluator credibility were strongly correlated with teachers' perceptions of the usefulness of the feedback. Although most teachers agreed or strongly agreed that their evaluator was credible, teachers' perceptions of their evaluator's credibility were found to be significantly related to the usefulness of the feedback, and 68–83 percent of teachers rated statements related to evaluator credibility as important in their response to feedback.
- Where teachers have less favorable ratings of evaluator credibility, considering ways to build evaluators' knowledge of the content or subject being evaluated, knowledge of how students learn, knowledge of teaching practices, understanding of the curriculum being observed, and understanding of the established teacher evaluation system.

The findings may also warrant additional research to further test the model for performance feedback in teacher evaluation systems with larger samples and to examine the relationship between teachers' response to feedback and teachers' performance to confirm the structural equation model.

Limitations of the study

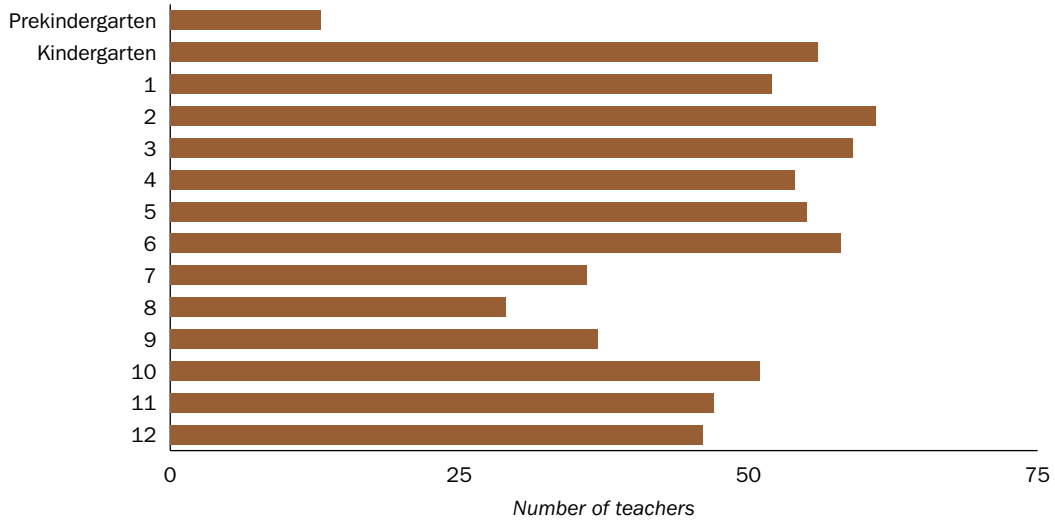
This study sought to include teacher performance as a variable in the model, but because of data limitations and variability, teacher performance was not examined. At the time of publication, some of the districts included in the study were not computing final performance ratings because they were still pilot-testing the inclusion of student growth as one measure of teacher performance, and thus they could not share final teacher performance ratings. Data from districts that had final performance ratings—that is, ratings that included both measures of teacher practice and student growth data—showed limited variability, with most teachers rated effective or highly effective, which made it difficult for the study team to examine relationships with performance. To address this limitation, future research could examine teacher ratings as states and districts implement new teacher evaluation systems with ratings that better capture differences in teacher performance. These ratings could be examined to determine how the characteristics of feedback, teachers' access to resources, and teachers' responses to feedback relate to teachers' performance.

Another limitation to the study is that it used a purposive sample with voluntary participation, so the findings may not be generalizable to other settings. It is possible that teachers who volunteered to participate in the study may differ in their responses from teachers who did not participate.

Appendix A. Analysis sample

Seven school districts in two Regional Educational Laboratory Central Region states participated in this study. These districts represented urban locales (44 percent), rural locales (31 percent), and small towns (25 percent). In each district a district contact provided teachers with a link to a video describing the study and a hard copy of the consent form. Teachers were asked to sign and return the consent form to the study team if they were interested in participating. Only teachers who were participating in the district's new teacher evaluation system and had direct student contact in a classroom setting were asked to participate. Of the 317 teachers who consented to the survey, 243 completed the survey, which is a response rate of 76.7 percent. Teachers who completed the survey represented all grade levels (figure A1) and a variety of subject areas (figure A2).

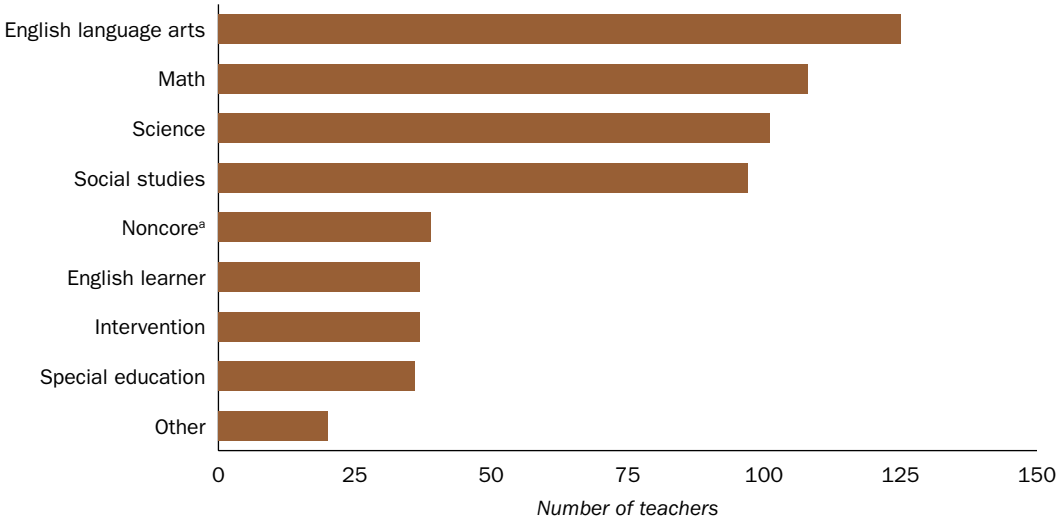
Figure A1. Teachers who completed the survey represented all grade levels, 2014/15



Note: Values do not sum to 243 because some teachers taught more than one grade.

Source: Authors' compilation.

Figure A2. Teachers who completed the survey represented a variety of subject areas, 2014/15



Note: Values do not sum to 243 because some teachers taught more than one subject.

a. Includes physical education, art, and technology.

Source: Authors' compilation.

Appendix B. Methods

This appendix describes the methods used in the study, including survey development, descriptive analysis, correlational analysis, and structural equation modeling.

Survey development

The Examining Evaluator Feedback Survey (Cherasaro et al., 2015) was developed as part of this study using an iterative process that included expert review, cognitive interviews, and statistical modeling (Presser et al., 2004; Rothgeb, 2008; see table B1 for descriptions of the questions in each section of the survey). A variety of statistical techniques (classical test theory, Rasch analysis, and confirmatory factor analysis) were used to examine the reliability and validity of the survey. These analyses provided evidence for the reliability and validity of the domains of usefulness, accuracy, evaluator credibility, access to resources, and response to feedback to inform state and district education leaders about teachers' perceptions of feedback from evaluators. The analyses also provided evidence of reliability of questions related to the importance of feedback characteristics.

Descriptive analysis

After teachers' responses were gathered from SurveyGizmo, the online survey administration platform used in this study, the study team cleaned the data to include only teachers who had received feedback for the 2014/15 school year and to remove duplicate responses.

Table B1. Questions by section in the Examining Evaluator Feedback Survey, 2015

Section	Description	Question number
Background information	Definition of designated evaluator	1
	Designated teacher evaluator in the current school year	2
	Frequency of feedback conversation with designated evaluator	3
	Frequency of written feedback from designated evaluator	4
Feedback characteristics (Includes five categories of questions)	Usefulness: perceived usefulness of evaluator's feedback	5 (a–g)
	Accuracy: perceived accuracy of evaluator's feedback	6 (a–d)
	Credibility: perceived credibility of evaluator	7 (a–e)
	Access to resources: perceived access to professional development and other resources needed to respond to evaluation feedback	8 (a–d)
	Responsiveness: actions teacher took in response to evaluation feedback	9 (a–e)
Importance of feedback characteristics	Importance of the following characteristics when deciding how to respond to feedback:	
	• Perceiving the feedback as useful	10 (a–i)
	• Having confidence in the accuracy of the feedback	11 (a–b)
	• Perceiving the evaluator as credible	12 (a–e)
	• Having access to relevant resources	13 (a–d)
Belief about instructional improvement	Belief about whether feedback improved teacher's instruction	14
Teacher demographics	Number of years teaching	15
	Grade level or levels currently teaching	16
	Subject area or areas currently teaching	17

Source: Cherasaro et al., 2015.

The study team then used SPSS software to calculate response frequencies for both the agreement questions and the importance questions.

Correlational analysis

Bivariate correlations and structural equation modeling were used to examine the relationship of usefulness, accuracy, evaluator credibility, access to resources, and response to feedback. First, a scale score was created for each domain by averaging the set of questions within the domain. Bivariate correlations between the scale scores for each domain were examined. All survey measures had high internal consistency (table B2). Bivariate correlations among teachers' ratings of usefulness, accuracy, and evaluator credibility showed these to be highly related to each other (see table 2 in the main text).

Structural equation modeling

To more thoroughly assess the influence of performance feedback on teachers' response to feedback, structural equation modeling, using maximum likelihood estimation, was used to test the hypothesized direction of influence. The hypothesized model was reviewed using a two-step process. First, the underlying measurement model was tested, by examining the entire set of measurement items simultaneously, to determine whether the survey items were good measures of their respective domains. Next, the structural model was tested to determine the relationship among the domains.

Measurement model

All survey items had high factor loadings on their respective domain (for example, all usefulness items had a strong positive relationship to the usefulness domain they composed; table B3). To determine model fit in structural equation modeling, multiple fit indexes should be considered. A review of several fit indexes suggests that the measurement model adequately fit the data (table B4). While the root mean square error of approximation value was greater than the target value for a good fit, the value does indicate an adequate fit based on the rule of thumb that values of less than 1 demonstrate reasonable fit (Kline, 2005). These findings replicate those found during pilot-testing of the survey.

Structural model

Structural equation modeling was used to identify a model that was the most parsimonious and best fit the data by first examining the fit of the model and the path coefficients. Prior to assessing the fit of the final model, the study team tested the fit of the data to the

Table B2. Reliability of survey domains

Survey domain	Cronbach s alpha
Usefulness	0.93
Accuracy	0.89
Evaluator credibility	0.92
Access to resources	0.80
Response to feedback	0.89

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Table B3. Survey item loadings by domain

Domain	Factor loading	Standard error
Usefulness		
Usefulness item 1	0.87	0.02
Usefulness item 2	0.83	0.02
Usefulness item 3	0.87	0.02
Usefulness item 4	0.82	0.02
Usefulness item 5	0.76	0.03
Usefulness item 6	0.73	0.03
Usefulness item 7	0.78	0.03
Accuracy		
Accuracy item 1	0.86	0.02
Accuracy item 2	0.85	0.02
Accuracy item 3	0.79	0.03
Accuracy item 4	0.79	0.03
Evaluator credibility		
Credibility item 1	0.88	0.02
Credibility item 2	0.85	0.02
Credibility item 3	0.87	0.02
Credibility item 4	0.87	0.02
Credibility item 5	0.73	0.03
Access to resources		
Access item 1	0.81	0.04
Access item 2	0.67	0.05
Access item 3	0.70	0.04
Access item 4	0.67	0.05
Response to feedback		
Response item 1	0.85	0.03
Response item 2	0.81	0.03
Response item 3	0.82	0.03
Response item 4	0.80	0.03
Response item 5	0.66	0.04

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Table B4. Measurement model fit

Statistic	Obtained value	Target value
Root mean square error of approximation (90 percent confidence interval)	0.09 (0.08 to 0.09)	≤0.08
Comparative fit index	0.90	>0.90
Tucker-Lewis index	0.88	>0.90
Standard root mean square residual	0.06	<0.10

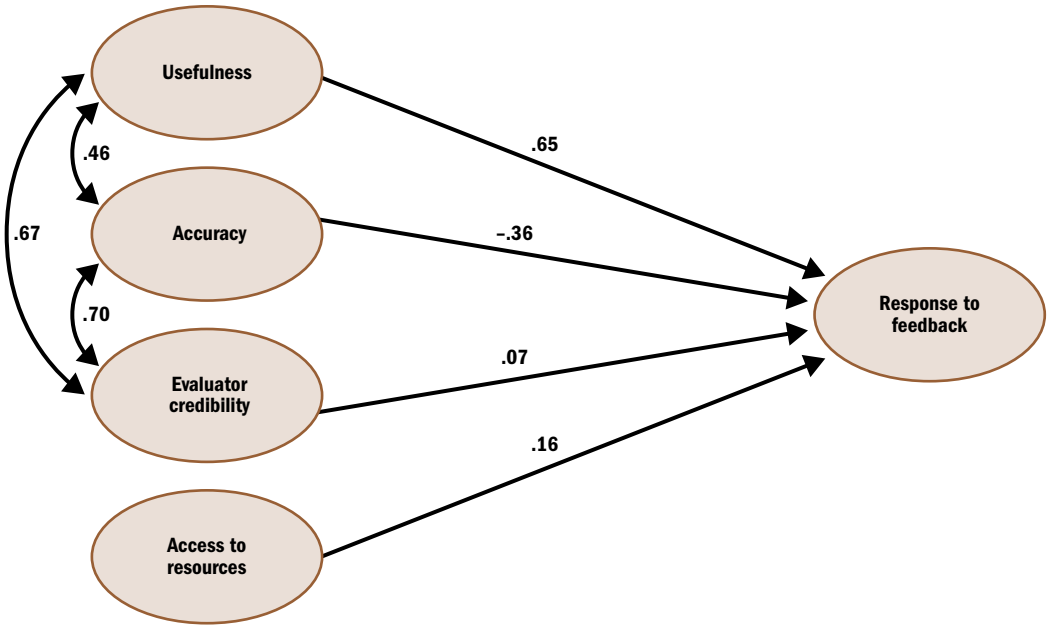
Note: Target values are approximate values needed to demonstrate good model fit (Kline, 2005).

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

hypothesized model. The hypothesized model did not result in any significant path coefficients and found a weak relationship between evaluator credibility and response to feedback and a negative relationship between accuracy and response to feedback (figure B1). In assessing a structural equation model, it is important to examine the model statistics and to examine the model with respect to substantive theory. The adjustments to identify a better fitting model should be in alignment with the theory that the model is intended to test (Hooper, Coughlan, & Mullen, 2008).

To determine adjustments to the model, the study team examined the correlational analysis, path coefficients, and modification indexes for the hypothesized model and considered changes that made theoretical sense. None of the modification indexes was large enough to justify adjustments to the model. The test of the hypothesized model and the correlational analysis suggested that accuracy and evaluator credibility have a strong correlation and that they are both correlated with usefulness, which had a strong positive relationship with response to feedback. Theoretically, teachers may need to believe that the feedback that they receive is accurate and that their evaluator is credible before they are able to find the feedback useful. Additionally, whether teachers believe that their feedback is accurate may be strongly related to the extent to which they believe that their evaluator is credible. The hypothesized model was adjusted so that usefulness, accuracy, and evaluator credibility were not three independent but related aspects of perceived feedback quality; rather, usefulness is determined partially by accuracy and evaluator credibility, resulting in the final model (see figure 10 in the main text). Although the final structural model did not fit

Figure B1. No significant path coefficients were found in the hypothesized model



Note: There were no significant relationships found in the hypothesized model. Arrows represent a relationship between variables, with the direction of the arrow showing which variable influenced the other. The numbers above the arrows are STDYX path coefficients, which are regression coefficients that are standardized on both the independent and dependent variables. Each number represents the standardized amount of change in the dependent variable that would occur with one standardized standard deviation in the independent variable—or the strength, direction, and significance of the relationship. Larger numbers indicate stronger relationships, and positive numbers indicate that an increase in one variable is related to an increase in the other variable.

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

Table B5. Final model fit

Statistic	Obtained value	Target value
Root mean square error of approximation	0.10	≤0.08
Comparative fit index	0.86	>0.90
Tucker-Lewis index	0.84	>0.90
Standard root mean square residual	0.20	<0.10

Note: Target values are approximate values needed to demonstrate good model fit (Kline, 2005).

Source: Authors' calculations based on responses to the Examining Evaluator Feedback Survey of 2015.

the data as well as desired (table B5), the pattern of relationships between the variables was the most parsimonious and interpretable of the alternative models considered.[3]

Notes

1. ESEA flexibility waivers allow states to request flexibility regarding the requirements of the No Child Left Behind Act of 2001 by proposing state-developed plans to address four principles: college- and career-ready expectations for all students; state-developed differentiated recognition, accountability, and support; support for effective instruction and leadership; and reduction in duplication and unnecessary burden.
2. Montana is the only state that has not requested a waiver. Some states requested waivers but later withdrew or revoked them.
3. Models in which accuracy and evaluator credibility were combined into one domain and in which usefulness, accuracy, evaluator credibility, and access to resources all covaried were also considered.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public schools. *Journal of Labor Economics*, *25*(1), 95–135. http://eric.ed.gov/?id=ED505650

Center on Education Policy. (n.d.). *Federal education programs/NCLB/ESEA waivers: Waiver watch*. Washington, DC: Author. Retrieved December 22, 2015, http://www.cep-dc.org/index.cfm?DocumentSubTopicID=48#Tracking.

Cherasaro, T. L., Brodersen, R. M., Yanoski, D. C., Welp, L. C., & Reale, M. L. (2015). *The Examining Evaluator Feedback Survey* (REL 2016–100). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. http://eric.ed.gov/?id=ED561139

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers 1: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632.

Coggshall, J. G., Rasmussen, C., Colton, A., Milton, J., & Jacques, C. (2012). *Generating teaching effectiveness: The role of job-embedded professional learning in teacher evaluation: A Research & Policy Brief*. Washington, DC: National Comprehensive Center for Teacher Quality. http://eric.ed.gov/?id=ED532776

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81–112. http://eric.ed.gov/?id=EJ658467

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). *Reducing behavior problems in the elementary school classroom: A practice guide* (NCEE No. 2008–012). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. http://eric.ed.gov/?id=ED502720

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

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Analysis and Policy Analysis*, *22*(2), 129–145. http://eric.ed.gov/?id=EJ615883

Hemmeter, M. L., Snyder, P., Kinder, K., & Artman, K. (2011). Impact of performance feedback delivered via electronic mail on preschool teachers' use of descriptive praise. *Early Childhood Research Quarterly*, *26*(1), 96–109. http://eric.ed.gov/?id=EJ906820

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, 6(1), 53–60.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), 349–371.

Johnson, K. A. (2000). *The effects of advanced teacher training in education on student achievement* (Center for Data Analysis Report No. 00–09). Washington, DC: Heritage Foundation. Retrieved June 21, 2012, from http://www.heritage.org/Research/Education/CDA00–09.cfm.

Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Methods of Effective Teaching, Bill & Melinda Gates Foundation. http://eric.ed.gov/?id=ED540961

Kinicki, A. J., Prussia, G. E., Wu, B., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *Journal of Applied Psychology*, 89(6), 1057–1069.

Kline, R. B. (2005). *Principles and practice of structural equation modeling, 2nd ed*. New York: Guilford Press.

Konstantopoulos, S., & Chung, V. (2010). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, *48*(2), 361–386. http://eric.ed.gov/?id=EJ921703

Monk, D. H., & King, J. (1994). Multi-level teacher resource effects on pupil performance in secondary mathematics and science: The role of teacher subject matter preparation. In R. Ehrenberg (Ed.), *Choices and consequences: Contemporary policy issues in education* (pp. 29–58). Ithaca, NY: ILR Press.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*(3), 237–257. http://eric.ed.gov/?id=EJ727526

Parise, L. M., & Spillane, J. P. (2010). Teaching learning and instructional change: How formal and on-the-job learning opportunities predict change in elementary school teachers' practice. *The Elementary School Journal*, *110*(3), 323–346. http://eric.ed.gov/?id=EJ890679

Presser, S., Rothgeb, J. M., Couper, M. P., Lessler, J. T., Martine, E., Martin, J., et al. (2004). *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: Wiley.

Rothgeb, J. M. (2008). Pilot test. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 583–585). Thousand Oaks, CA: Sage.

Scheeler, M. C., Ruhl, K. L., & McAfee, J. K. (2004). Providing performance feedback to teachers: A review. *Teacher Education and Special Education*, *27*(4), 396–407. http://eric.ed.gov/?id=EJ694123

Supovitz, J. (2012, April). *The linking study—First year results: A report of the first year effects of an experimental study of the impact of feedback to teachers on teaching and learning.* Paper presented at the 2012 Annual Meeting of the American Educational Research Association, Vancouver, BC. Retrieved June 21, 2012, from http://www.cpre.org/linking-study-first-year-results.

Tuytens, M., & Devos, G. (2011). Stimulating professional learning through teacher evaluation: An impossible task for the school leader? *Teaching and Teacher Education*, *27*(5), 891–899. http://eric.ed.gov/?id=EJ924136

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*(1), 57–67. http://eric.ed.gov/?id=EJ548364

U.S. Department of Education. (2012). *ESEA flexibility: State requests and related documents*. Washington, DC: Author. Retrieved June 21, 2012, from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html.

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