Tools

November 2015

The Examining Evaluator Feedback Survey

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

Overview

The Examining Evaluator Feedback Survey is a tool for administrators to gather information on teachers' perceptions of the feedback they receive from their evaluator and on teachers' self-reported responses to that feedback. This report contains the survey questions and information about how the survey was developed and how it can be used.

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
Arne Duncan, *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 2016–100

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 2015

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.

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. 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

Regional Educational Laboratory Central, together with members of the Educator Effectiveness Research Alliance, developed the Examining Evaluator Feedback Survey to help administrators gather information from teachers about their perceptions of evaluator feedback and teachers' self-reported responses to that feedback. Evaluator feedback is defined as feedback on teaching performance that teachers receive from a designated evaluator as part of a formal district evaluation process.

District and state administrators can use this survey to systematically collect information on teacher perceptions on five key characteristics of evaluator feedback:

- Usefulness.
- Accuracy.
- Credibility.
- Access to resources.
- Responsiveness.

Administrators can also use the information gathered from the survey to better understand teacher perceptions of new evaluator feedback protocols.

This report presents the Examining Evaluator Feedback Survey along with information about how it was developed, how it can be used, and its reliability and validity.

Contents

Introduction

The Examining Evaluator Feedback Survey is designed to gather information from teachers about their perceptions of evaluator feedback and teachers' self-reported responses to that feedback. District and state administrators can use this survey to systematically collect information on teacher perceptions on five key characteristics of evaluator feedback:

- Usefulness.
- Accuracy.
- Credibility.
- Access to resources.
- Responsiveness.

The survey (provided in appendix A) takes 10–15 minutes to complete.

Why and how was this survey developed?

As new teacher evaluation systems that focus on teacher development have been developed, state and district administrators have expressed a growing interest in learning more about the quality and usefulness of the feedback provided to teachers

As of 2014, 45 states have requested Elementary and Secondary Education Act flexibility waivers that include plans to improve the utility of their teacher evaluation systems by providing targeted and ongoing feedback that informs teachers about their practice (U.S. Department of Education, 2014). While only five of seven states in the Regional Educational Laboratory (REL) Central Region requested a flexibility waiver (Colorado, Kansas, Missouri, South Dakota, Wyoming[1]), all are developing or implementing new teacher evaluation systems that focus on teacher development. As these systems have been developed, state and district administrators have expressed a growing interest in learning more about the quality and usefulness of the feedback provided to teachers.

In response to needs identified by members of the Educator Effectiveness Research Alliance, REL Central conducted a study during the 2014/15 school year to examine relationships among feedback characteristics (perceived utility and accuracy of evaluator feedback and credibility of the person providing the feedback), access to resources related to the feedback, teacher response to feedback, and teacher performance. The Examining Evaluator Feedback Survey was developed to collect information for that study.[2] The theoretical framework and literature used to develop the survey are in appendix B.

The survey was designed using an iterative process that included researchers and practitioners involved in the Educator Effectiveness Research Alliance. Early iterations of the survey focused on question development and cognitive pretesting, and later refinements focused on increasing the reliability of the data and enhancing the format of the instrument (see appendix C for additional details).

Researchers examined the reliability and validity of the survey using a variety of statistical techniques (classical test theory, Rasch analysis, and confirmatory factor analysis; see appendix C for a more complete explanation). These analyses provided evidence for the reliability and validity of the questions and the categories of usefulness, accuracy, credibility, access to resources, and responsiveness to inform state and district leaders about teachers' perceptions of evaluator feedback. The analysis also provided evidence of reliability of questions related to the importance of feedback characteristics.

Why administer this survey?

The Examining Evaluator Feedback Survey is a tool for district and state leaders and administrators to learn more about teachers' perceptions of the feedback they received as a part of their evaluation system. District and state leaders can use the results of this survey to inform decisions regarding feedback in teacher evaluation systems. For example, low ratings of feedback accuracy or evaluator credibility may suggest areas for improvement or topics for which additional training for evaluators could be provided. Teacher perceptions of the utility of feedback may suggest changes to evaluation policies and procedures such as timeliness and frequency and a focus on the types of feedback that teachers identify as most important. In addition, the survey can provide information about how teachers use the feedback that they receive, allowing administrators to consider ways to tailor professional development and provide resources for teachers to maximize professional growth.

The survey can provide information about how teachers use the feedback that they receive, allowing administrators to consider ways to tailor professional development and provide resources for teachers to maximize professional growth

The survey is not recommended for use in high-stakes decisionmaking, such as personnel or program funding decisions. The survey is intended only to gather teachers' perceptions to identify ways to support evaluators in providing feedback. When making high-stakes decisions, multiple data sources should be used that include more observational data and artifacts rather than perceptual data.

The survey questions

The questions in the Examining Evaluator Feedback Survey ask respondents for information in five sections: background information, five feedback characteristics (usefulness, accuracy, credibility, access to resources, and responsiveness), importance of feedback characteristics, belief about instructional improvement, and teacher demographics (table 1). It asks teachers to indicate their level of agreement or disagreement with a series of statements on a five-point scale and to indicate how important various aspects of evaluator feedback are to them.

How to administer and adapt this survey

You can adjust the survey to fit your specific needs and can distribute the survey to teachers in a pencil-and-paper format or using online software of your choosing. The survey takes approximately 10–15 minutes to complete. You should administer the survey as close to the end of the school year as possible, because teachers are asked to reflect on the feedback that they received throughout the entire school year. If possible, teacher responses should be collected anonymously so that teachers will respond more honestly, without concern that their administrator might see their individual responses. The survey should also be administered only to classroom teachers because the survey asks about changes in classroom management, instructional practice, and content or subject knowledge. The survey questions are not relevant to teachers who work with students individually or who work only with other teachers to provide support.

You are free to adapt any part of the survey for your own use. For example, you may be interested in collecting data for only a subset of the categories targeted by the survey. If you decide to use a subset of the questions within a category or to adapt the language of the questions or response options for your own use, you are encouraged to establish reliability and validity for the revisions (see appendix C for statistical techniques that you could use to do so).

Table 1. Questions in the Examining Evaluator Feedback Survey, by section

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 evaluator feedback	8 (a–d)
	Responsiveness: actions teacher took in response to evaluator feedback	9 (a–e)
Importance of feedback characteristics	Importance of the following characteristics when deciding how to respond to the feedback: • Perceiving the feedback as useful. • Having confidence in the accuracy of the evaluation feedback. • Perceiving the evaluator as credible. • Having access to relevant resources.	 10 (a–i) 11 (a–b) 12 (a–e) 13 (a–d)
Belief about instructional improvement	Belief about whether evaluator 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: Authors' compilation based on the survey in appendix A.

How to analyze survey results

There are four steps involved in analyzing the data collected using the Examining Evaluator Feedback Survey.

Step 1. Examining response rates

This step involves reviewing the collected data to see whether there are any significant gaps. For example, if only 50 percent of the surveyed teachers respond to a given question, the responses for that question may not be useful or representative of the group. You may also consider why data are missing; the fact that teachers did not respond to a question may be as revealing as the responses themselves.

Step 2. Determining what questions will be answered using the data

Table 2 lists basic research questions that can be answered by the survey data, and the analysis needed for each question.

When analyzing responses to questions in the five categories in the feedback characteristics section, you may choose to look at aggregate responses to individual questions or at all questions in the category as a whole. However, responses to the groups of questions in the importance of feedback characteristics section should be examined by individual question only.[3]

Table 2. Sample research questions that could be answered using the data, along with possible analysis methods to answer them and sample findings

Research question	Possible analysis method	Sample question	Sample finding (examples, not the result of any analysis that was conducted)
How did teachers respond in aggregate to an individual question?	Percentage or median response to individual questions	How did all teachers respond to the question "My evaluator's feedback included specific performance suggestions"?	46 of 50 teachers (92%) agreed or strongly agreed that the evaluator's feedback included specific performance suggestions
How did teachers respond in aggregate to each of the categories?	Aggregate percentage or median response across an entire category	How did all teachers respond to all the questions in the usefulness category?	26 of 50 teachers (52%) agreed or strongly agreed that they received useful feedback
How did subgroups of teachers respond in aggregate to an individual question?	Responses to a question by groups of teachers using background or teacher demographic questions	How did grade 6 teachers respond to the question "My evaluator's feedback included specific performance suggestions"?	10 of 15 grade 6 teachers (67%) agreed or strongly agreed that the evaluator's feedback included specific performance suggestions
How did subgroups of teachers respond in aggregate to all the questions in a category?	Responses for a category disaggregated by background question or teacher demographic	How did grade 6 teachers respond to all the questions in the usefulness category?	12 of 15 grade 6 teachers (80%) agreed or strongly agreed that they received useful feedback

Source: Authors' compilation.

Step 3. Producing reports based on the questions that were answered

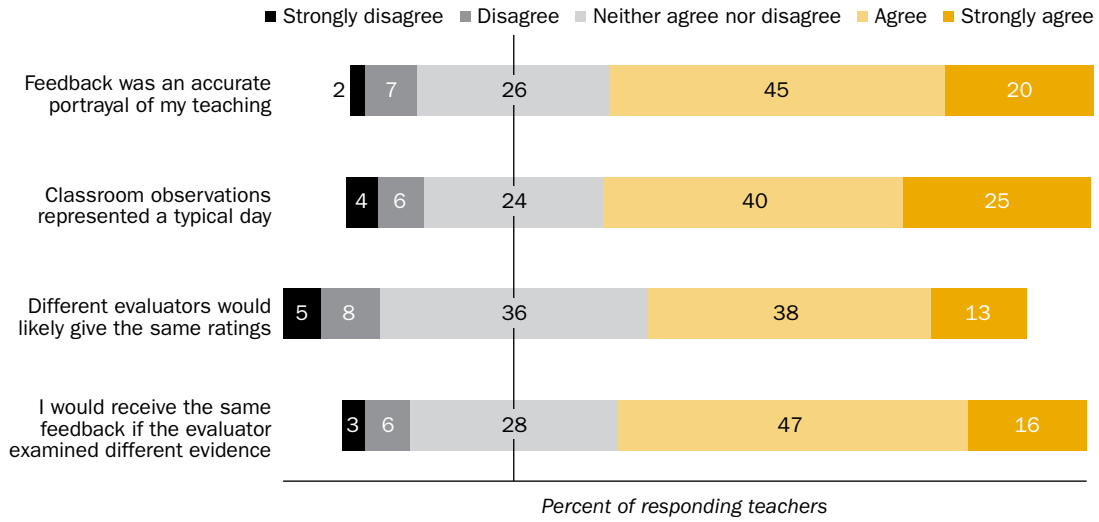
If the survey is administered online, you could use the results to create reports with graphs that display results across a category, by individual question, or by teacher subgroup. If the survey is administered in a paper-and-pencil format, you could input the responses into a spreadsheet and use formulas to calculate the percentage of responding teachers or the median.

An analysis using sample data of percentage of teachers responding for individual questions is presented in figure 1, and an analysis using sample data of percentage of responding teachers by category is presented in figure 2.

Step 4. Interpreting the results

This step involves investigating patterns in the data to answer the research questions identified in step 2. For example, you may want to know whether the feedback that was provided to teachers was useful to them. To determine this, you might examine the findings from figure 2 and discover that almost half the responding teachers did not agree that the feedback was useful: 36 percent neither agreed nor disagreed, and 13 percent disagreed or strongly disagreed. At this point, it might be helpful to re-examine the data and look at the individual questions within a category. Each question asks about a different component of the category, so looking at the response to individual questions could allow you to more precisely determine respondents' needs. It also might be useful to collect additional data to help interpret the findings—for example, through follow-up interviews or focus groups—to provide context for the survey results.

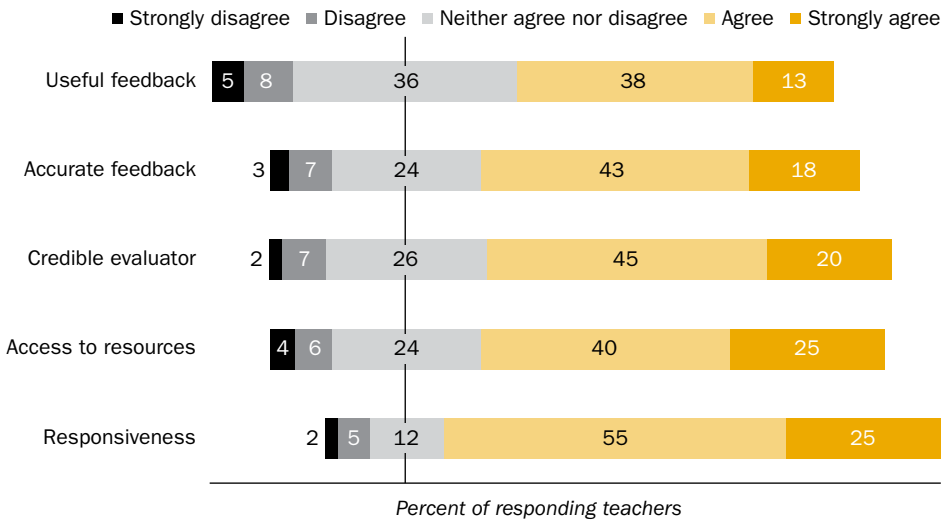
Figure 1. Sample analysis of responses to accuracy questions



Note: The vertical bar represents the median point of the neutral response category.

Source: Authors' compilation.

Figure 2. Sample analysis of responses by category



Note: The vertical bar represents the median point of the neutral response category.

Source: Authors' compilation.

In this example, examining the individual questions within the category might have revealed that the feedback was not useful because it did not provide suggestions, strategies, or resources and was not timely. Given those results, you might focus additional evaluator training on providing feedback with specific suggestions for strategies or professional development or consider how to address barriers to the frequency and timeliness of feedback.

A variety of responses are available to address the needs revealed by the data analysis (table 3).

Table 3. How to respond to needs in different feedback characteristic categories

Feedback characteristic categories	Possible responses to needs in this category
Usefulness	• Providing additional training for evaluators. • Implementing system changes to increase timeliness of feedback. • Increasing the amount and frequency of feedback provided. • Including recommendations. • Implementing a school- or district-wide instructional practice.
Accuracy	• Providing additional training for evaluators. • Providing training to improve inter-rater reliability. • Increasing the number of observations. • Modifying or improving evaluation tools or rubrics. • Implementing a school- or district-wide instructional practice.
Credibility	• Providing additional training for evaluators. • Using subject matter experts as evaluators. • Using peer evaluators. • Implementing a school-or district-wide instructional practice. • Increasing the number of observations.
Access to resources	• Increasing or improving available district professional development resources. • Scheduling changes to allow for job-embedded professional development. • Implementing peer or mentor coaching models. • Establishing partnerships with local education agencies or postsecondary institutions. • Using evaluation data to determine the focus of professional development.
Responsiveness	• Implementing peer or mentor coaching models. • Including action steps as a component of feedback. • Increasing the number of observations. • Taking personnel actions. • Implementing peer or mentor coaching models.

Source: Authors' compilation.

* * *

The Examining Evaluator Feedback Survey is a tool that administrators can use to gather data to inform decisionmaking regarding feedback in teacher evaluation systems. The survey (provided in appendix A) collects teacher perceptions on five key aspects of evaluator feedback: usefulness, accuracy, credibility, access to resources, and responsiveness. The survey was developed based on the current literature on performance feedback (see appendix B) and was reviewed and tested for evidence of reliability and validity (see appendix C).

Appendix A. The Examining Evaluator Feedback Survey

The purpose of this survey is to understand your thoughts on the usefulness and accuracy of your evaluator feedback. The survey asks questions about your experiences with the feedback you received as part of your district's teacher evaluation system. As you answer the questions, please consider only feedback that you received from your designated evaluator in your district during the current school year. Your designated evaluator is the person who is responsible for providing your performance rating at the end of the school year.

1. I have read and understand these instructions.

   ☐ Yes

2. As part of the district's teacher evaluation system, who was your designated evaluator in the current school year? (Select only one. If you have more than one evaluator please pick one and refer to that evaluator as you respond to the remaining questions.)

   ☐ My principal
   ☐ My assistant principal
   ☐ A peer
   ☐ My department chair
   ☐ My coach
   ☐ Other (please describe): _____

3. How often did you have a feedback conversation with your designated evaluator throughout the current school year? Feedback conversations are defined as any conversation with your evaluator in which he or she provided feedback specific to observations, walkthroughs, or artifacts collected as part of your evaluation.

   ☐ Never
   ☐ Once
   ☐ Twice
   ☐ Three times
   ☐ Four times
   ☐ Five times
   ☐ More than five times

4. How often did you receive written feedback from your designated evaluator throughout the current school year? Written feedback includes feedback specific to observations, walkthroughs, or artifacts collected as part of your evaluation that was given to you in written form (either on paper or electronically).

   ☐ Never
   ☐ Once
   ☐ Twice
   ☐ Three times
   ☐ Four times
   ☐ Five times
   ☐ More than five times

For the following questions please keep in mind the feedback that you received throughout the current school year from your designated evaluator.

5. Indicate your level of agreement with the following statements. My evaluator's feedback…

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
… included specific improvement suggestions.	☐	☐	☐	☐	☐
… included specific suggestions to improve my content/ subject knowledge.	☐	☐	☐	☐	☐
… included specific instructional strategies that I could use to improve my teaching.	☐	☐	☐	☐	☐
… included specific classroom management strategies that I could use to improve my teaching.	☐	☐	☐	☐	☐
… included recommendations for finding resources or professional development to improve my teaching.	☐	☐	☐	☐	☐
… was provided as frequently as I needed it.	☐	☐	☐	☐	☐
… was provided in time for me to use it to inform my practice.	☐	☐	☐	☐	☐

6. Indicate your level of agreement with the following statements.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The feedback I received was an accurate portrayal of my teaching.	☐	☐	☐	☐	☐
The classroom observations or walkthroughs that informed the feedback I received represented a typical day in my classroom.	☐	☐	☐	☐	☐
The evaluation system is accurate enough that different evaluators reviewing the same evidence would likely give the same ratings.	☐	☐	☐	☐	☐
I would receive the same feedback if my evaluator examined different evidence (e.g., if they observed additional lessons or reviewed additional evidence).	☐	☐	☐	☐	☐

7. Indicate your level of agreement with the following statements. In my opinion, my evaluator had sufficient …

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
… knowledge of my content/subject to effectively evaluate me.	☐	☐	☐	☐	☐
… knowledge of how my students learn to effectively evaluate me.	☐	☐	☐	☐	☐
… knowledge of effective teaching practices to effectively evaluate me.	☐	☐	☐	☐	☐
… understanding of the curriculum being observed to effectively evaluate me.	☐	☐	☐	☐	☐
… understanding of the established teacher evaluation system to effectively evaluate me.	☐	☐	☐	☐	☐

8. Indicate your level of agreement with the following statements.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I had access to the professional development (formal or informal) that I needed in order to implement suggestions provided in my feedback.	☐	☐	☐	☐	☐
I had access to an instructional leader (e.g., peer, coach/ mentor, administrator) who supported me in implementing suggestions provided in my feedback.	☐	☐	☐	☐	☐
I was able to observe expert teachers modeling skills that related to my feedback.	☐	☐	☐	☐	☐
I had time during the school day to plan for implementing new strategies based on my feedback (e.g., collaborative or individual planning time).	☐	☐	☐	☐	☐

9. Indicate your level of agreement with the following statements. Because of the feedback I received from my evaluator …

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
… I tried new instructional strategies in my classroom.	☐	☐	☐	☐	☐
… I tried new classroom management strategies in my classroom.	☐	☐	☐	☐	☐
… I sought professional development opportunities (formal or informal).	☐	☐	☐	☐	☐
… I sought advice from an instructional leader (for example, peer, coach or mentor, administrator).	☐	☐	☐	☐	☐
… I changed the way I plan instruction.	☐	☐	☐	☐	☐

10. When deciding how to respond to your feedback, how important was each the following? Receiving …

	Unimportant	Slightly Important	Important	Very Important	Critical
… specific improvement suggestions.	☐	☐	☐	☐	☐
… recommended next steps for finding professional development to improve your teaching.	☐	☐	☐	☐	☐
… feedback within an appropriate timeframe.	☐	☐	☐	☐	☐
… feedback as frequently as you needed it.	☐	☐	☐	☐	☐
… feedback with specific suggestions to improve your content or subject knowledge.	☐	☐	☐	☐	☐
… specific instructional strategies that you could use to improve your teaching.	☐	☐	☐	☐	☐
… specific classroom management strategies that you could use to improve your teaching.	☐	☐	☐	☐	☐
… feedback that was an accurate portrayal of my teaching.	☐	☐	☐	☐	☐
… feedback from classroom observations or walkthroughs that represented a typical day in my classroom.	☐	☐	☐	☐	☐

11. When deciding how to respond to your feedback, how important was each the following? Having confidence that I would receive the same feedback …

	Unimportant	Slightly Important	Important	Very Important	Critical
… from a different evaluator if they reviewed the same evidence.	☐	☐	☐	☐	☐
… if my evaluator had examined different evidence (e.g., if they observed additional lessons or reviewed additional evidence).	☐	☐	☐	☐	☐

12. When deciding how to respond to your feedback, how important was each the following? Having confidence that my evaluator had sufficient …

	Unimportant	Slightly Important	Important	Very Important	Critical
… knowledge of my content/subject to effectively evaluate me.	☐	☐	☐	☐	☐
… knowledge of how my students learn to effectively evaluate me.	☐	☐	☐	☐	☐
… knowledge of effective teaching practices to effectively evaluate me.	☐	☐	☐	☐	☐
… understanding of the curriculum being observed to effectively evaluate me.	☐	☐	☐	☐	☐
… understanding of the established teacher evaluation system to effectively evaluate me.	☐	☐	☐	☐	☐

13. When deciding how to respond to your feedback, how important was each the following?

	Unimportant	Slightly Important	Important	Very Important	Critical
Having access to the professional development (formal or informal) that I needed in order to implement suggestions provided in my feedback.	☐	☐	☐	☐	☐
Having access to an instructional leader (e.g., peer, coach/mentor, administrator) who supported me in implementing suggestions provided in my feedback.	☐	☐	☐	☐	☐
Being able to observe expert teachers modeling skills that related to my feedback.	☐	☐	☐	☐	☐
Having time during the school day to plan for implementing new strategies based on my feedback (e.g., collaborative or individual planning time).	☐	☐	☐	☐	☐

For the following question please keep in mind the feedback that you received throughout the current school year from your designated evaluator.

14. To what extent did the feedback you received from your designated evaluator improve your instruction?

- ☐ Not at all
- ☐ A little
- ☐ A lot

15. Including this year, please indicate how many years of teaching experience you have.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7
- ☐ 8
- ☐ 9
- ☐ 10
- ☐ 11
- ☐ 12
- ☐ 13
- ☐ 14
- ☐ 15
- ☐ 16
- ☐ 17
- ☐ 18
- ☐ 19
- ☐ 20
- ☐ More than 20

16. Please indicate the grade level that you teach currently (select one or more).

- ☐ Early childhood
- ☐ Kindergarten
- ☐ Grade 1
- ☐ Grade 2
- ☐ Grade 3
- ☐ Grade 4
- ☐ Grade 5
- ☐ Grade 6
- ☐ Grade 7
- ☐ Grade 8
- ☐ Grade 9
- ☐ Grade 10
- ☐ Grade 11
- ☐ Grade 12

17. Please indicate the subject and students that you teach currently (select one or more).

- ☐ Language arts
- ☐ Math
- ☐ Science
- ☐ Social studies
- ☐ Noncore subjects (physical education, art, technology)
- ☐ English learner students
- ☐ Students in special education
- ☐ Intervention
- ☐ Other: _____

Appendix B. Related literature

The Examining Evaluator Feedback Survey was developed as part of Regional Educational Laboratory Central's study on teachers' reports of their experiences with feedback and how these perceptions influence their use of feedback and their performance. Specifically, the study examines relationships among five feedback characteristics: usefulness, accuracy, credibility, access to resources, responsiveness, and teacher performance. The survey was developed because no single survey instrument existed to assess all the feedback characteristics and categories of interest.

Ilgen, Fisher, and Taylor's (1979) mediation theory, used to guide development of the survey, describes a process through which feedback influences performance. The model suggests that performance is influenced by the use of feedback, which depends on the recipient's initial perception of the feedback, acceptance of the feedback, desire to respond to the feedback, intended response to the feedback, and external constraints. Initial perception of feedback "is concerned with how accurate the recipient perceives the feedback from any given source" (Ilgen et al., 1979, p. 353). This initial perception of feedback can be influenced by the source providing the feedback and by the timeliness and frequency of feedback. Once a recipient has formed an initial perception of the feedback, acceptance becomes an issue. Acceptance of feedback "refers to the recipient's belief that the feedback is an accurate portrayal of his or her performance" (Ilgen et al., 1979, p. 356). The acceptance of feedback can be influenced by the perception of the credibility of the source and the specificity of the feedback (especially when feedback is negative). The final stages of Ilgen et al.'s (1979) model are the desire to respond and intended response, described as a person's willingness or motivation to respond. In the model, Ilgen et al. (1979) also explain that external constraints, such as lack of resources or skills necessary to implement feedback, may influence the recipient's response.

Building on the Ilgen et al. (1979) model, recent studies have examined models of performance feedback to describe the relationship among variables that Ilgen et al. identified as important in the use of feedback to improve performance. Specifically, studies examine the relationship among the following feedback characteristics: usefulness of feedback, accuracy of feedback, credibility of the person providing feedback, and the effects of external constraints that may influence teacher response to feedback, such as access to resources related to feedback. Kinicki, Prussia, Wu, and McKee-Ryan (2004) studied relationships between supervisor feedback and loan officers' performance, and Tuytens and Devos (2011) studied relationships between supervisor feedback and teacher pursuit of professional learning in secondary schools. As predicted by Ilgen et al. (1979), both studies found that differences in feedback characteristics relate to an individual's responsiveness to the feedback and ultimately to future performance. Details of the findings from Kinicki et al. (2004) and Tuytens and Devos (2011), as well as emerging literature on the three feedback characteristics (usefulness, accuracy, and credibility) and access to learning opportunities as they relate to using feedback to improve performance, are described in detail below.

Usefulness of feedback

Usefulness of feedback depends on both the specificity of feedback and the timeliness and frequency of feedback. Kinicki et al. (2004) included the variable "feedback-rich environment" in their model of performance feedback. This variable included frequency of

feedback, specificity of feedback, and proportion of positive and negative feedback. They found that a feedback-rich environment was positively related to the perceived accuracy of feedback. Similarly, Tuytens and Devos (2011) found that the specificity and utility of feedback that a teacher had received during a supervisory evaluation conference had a direct relationship to that teacher's professional learning activities. Although neither of these studies prescribes an optimal level of specificity or frequency, both suggest that perceptions about the specificity and frequency influence the use of feedback.

Specific corrective feedback, which includes particular suggestions, directions, or examples of how to use an instructional strategy more appropriately and effectively, has been found to result in improved teaching performance compared with general feedback (such as "good" or "right"; Hemmeter, Snyder, Kinder, & Artman, 2011; Scheeler, Ruhl, & McAfee, 2004). In a recent and ongoing study of the effects of feedback for teachers in elementary school, teachers who were provided with specific feedback on their observed instructional practices had significantly greater instructional academic rigor, accountable talk, and higher student achievement than teachers who were provided with only student data as feedback (Supovitz, 2012). In addition to feedback on individual instructional strategies, specific feedback to increase content or subject knowledge and to improve classroom management strategies could theoretically lead to higher student achievement. Research on teacher quality and effective teaching suggests that subject matter knowledge (especially in math and science) influences student achievement (Goldhaber & Brewer, 2000; Monk & King, 1994; Johnson, 2000) and that implementing certain behavior management strategies can also lead to higher student engagement and achievement (Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008).

Timeliness of feedback may also be related to more effective use of feedback. Scheeler et al. (2004) identified and reviewed the results of 10 studies examining the effects of different feedback characteristics on teacher performance in pre-service teachers. The most consistent finding across the studies was that timeliness of feedback had a positive impact on teacher responsiveness.

Accuracy of feedback

Accuracy of feedback is the extent to which the person receiving feedback believes that the feedback accurately represents his or her performance. Kinicki et al. (2004) found that feedback that tends to be more specific, frequent, and positive was perceived as more accurate. Furthermore, perceived accuracy mediated the effect of these feedback features on intent to respond and performance (Kinicki et al., 2004).

Credibility of the person providing feedback

Credibility of the person providing feedback is the extent to which the person receiving feedback believes that the person providing the feedback is qualified to do so. Kinicki et al. (2004) found that perceived credibility of the source is related to both perceived accuracy and intent to respond. Additionally, participants who received more specific, frequent, and positive feedback perceived the source as more credible (Kinicki et al., 2004). Similarly, Tuytens and Devos (2011) found that teachers' perceived credibility of their principals as supervisors related to how they decided to use the feedback from their principals. Another important component of credibility is that the evaluator understands the evaluation

standards and has the ability to use the standards in feedback conversations (Coggshall, Rasmussen, Colton, Milton, & Jacques, 2012).

Access to resources

Theoretical and empirical research on teacher learning and professional growth also addresses the role of feedback. In particular, studies suggest that the use of a language of instruction, or a model of effective teaching and its decomposition, are seen as important to the development of expertise in teaching (Grossman et al., 2009). Increases in teachers' knowledge and skills and changes in their practice may be related to their access to resources that are aligned to their content area and specific needs, which may include allowing teachers to observe expert teachers, allowing teachers to engage in conversations with colleagues such as a coach or mentors about strategies, and helping teachers plan for implementation of new teaching approaches (Desimone, Porter, Garet, Yoon, & Birman, 2002; Garet, Porter, Desimone, Birman, & Yoon, 2001; Parise & Spillane, 2010).

Appendix C. Survey development methods

The survey was developed using an iterative process that included such survey development methods as expert review, cognitive interviews, and statistical modeling (Presser et al., 2004; Rothgeb, 2008). The original survey questions were formulated based on previous research (see appendix B) as well as unpublished surveys that the study team had implemented in evaluations of various teacher evaluation systems. The questions were reviewed and revised based on feedback from an advisory panel and teachers. The survey was then administered to 196 teachers, and the results were used to examine its reliability and validity. The stages of development are described in figure C1.

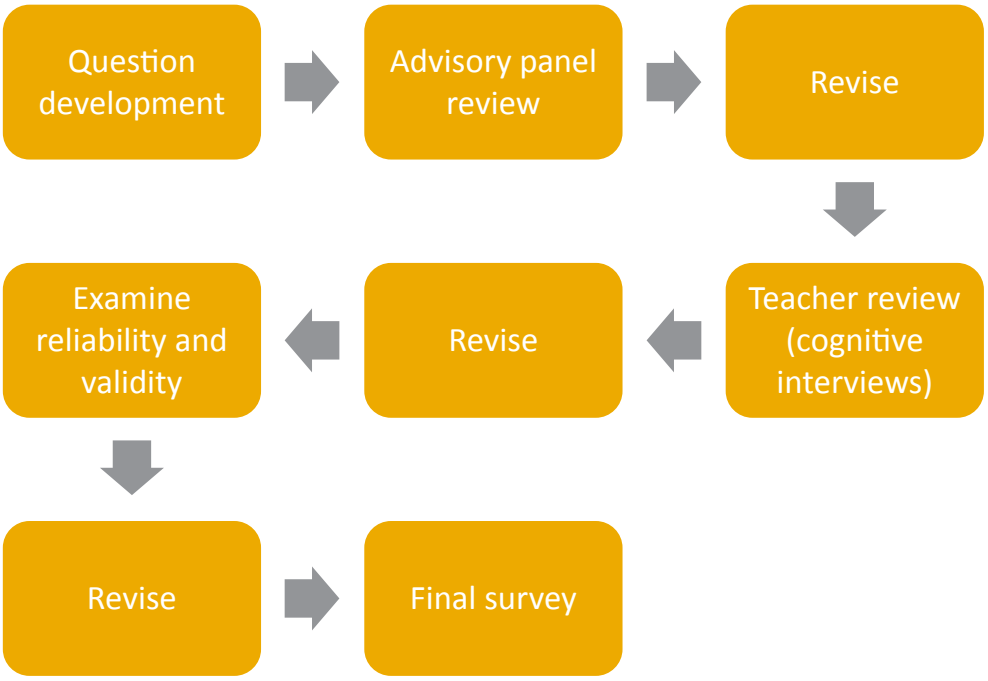
Advisory team review

To determine whether the survey was relevant to teacher evaluation systems in different contexts and whether the survey had face validity, the study team conducted a webinar with an advisory panel in which the panel responded to questions about the clarity and applicability of the questions and the appropriateness of the directions and response options. The advisory panel comprised seven members, including expert survey developers, state leaders, and district leaders with oversight for educator evaluation systems. Based on this review, the study team revised the directions, question stems, question wording, and response options on several questions for clarification and to increase ease in responding.

Cognitive interviews

To determine whether the survey questions were clear and would be uniformly interpreted by teachers, the study team conducted cognitive interviews with a sample of teachers. Nine

Figure C1. Survey development process



Source: Authors' illustration.

teachers responded to the survey and were interviewed by the study team using a structured interview protocol. Based on feedback from teachers, the study team made minor adjustments to the language of three questions.

Reliability and validity analysis

Following the cognitive interviews, the revised survey was administered to 196 teachers, 190 of whom completed the full survey. Survey reliability and validity were examined using classical test theory,[4] Rasch analysis,[5] and confirmatory factor analysis.[6] Analyses were conducted on the questions related to usefulness, accuracy, credibility, access to resources, and responsiveness. These categories showed high internal consistency, with Cronbach's alphas of 0.827–0.939. Confirmatory factor analysis suggested that the questions in these categories represent five distinct, though interrelated, categories (tables C1 and C2). Because the survey was developed as part of another study that is interested in analyzing descriptive statistics rather than testing a model for the importance and belief questions (10–14), Cronbach's alpha was examined for all items in those questions. The importance of feedback questions are numbered in question sets because they share the same question prompts. The prompts were used to make the survey easier to take. The importance of feedback questions showed high internal consistency, with a Cronbach's alpha of 0.931.

Rasch analysis was conducted on the category questions to examine whether respondents used the questions and the question response options as intended. Results indicated that the category questions successfully placed respondents along a continuum represented by the relevant category. However, Andrich threshold values[7] and probability curves suggested that respondents had difficulty distinguishing between the response options "somewhat disagree" and "somewhat agree." This finding was observed across all questions. Based on these findings, researchers decided to collapse these response options into a "neither agree nor disagree" option and form a five-point scale by recoding the survey data. All subsequent

Table C1. Examining Evaluator Feedback Survey factor loadings

Category	Factor loading range
Usefulness	0.23[a]–0.84
Accuracy	0.56–0.74
Credibility	0.59–0.86
Access to resources	0.37–0.66
Responsiveness	0.57–0.79

a. The item with the loading of 0.23 was revised; the remaining loadings ranged from 0.52 to 0.84.

Source: Authors' analysis based on pilot survey data.

Table C2. Examining Evaluator Feedback Survey factor standardized correlations

Category	Usefulness	Accuracy	Credibility	Resources
Accuracy	0.58			
Credibility	0.61	0.74		
Access to resources	0.72	0.66	0.65	
Responsiveness	0.59	0.023	0.31	0.56

Source: Authors' analysis based on pilot survey data.

reliability and validity analyses used this five-point response scale, and the scale was incorporated into the final survey.

The category minimum and maximum scores, mean, standard deviation, and reliability, after the two middle response options (somewhat disagree and somewhat agree) were combined to form a five-point response scale are presented in table C3. The table shows all categories to have an acceptable internal reliability, with respondents scoring along the full range of the response scale, except for responsiveness.

A confirmatory factor analysis using robust maximum likelihood estimation was also conducted to examine the structure of the survey. A measurement model was estimated where each scale question was an indicator of only its relevant category. Model fit indices suggested the model fit the data reasonably well. Factor loadings suggested the usefulness scale question "My evaluator's feedback was provided within an appropriate timeframe" be omitted or reworded. This question was reworded to "My evaluator's feedback was provided in time for me to use it to inform my practice."

Table C3. Examining Evaluator Feedback Survey scale descriptive statistics and reliabilities, by category

Category	*N*	Minimum	Maximum	Mean	Standard deviation	Cronbach s alpha
Usefulness	188	1.00	5.00	3.41	0.91	.929
Accuracy	187	1.00	5.00	3.65	0.84	.849
Credibility	187	1.00	5.00	3.75	0.94	.939
Access to resources	186	1.00	5.00	3.25	0.87	.824
Responsiveness	188	1.40	5.00	3.72	0.80	.917

Source: Authors' analysis based on pilot survey data.

Notes

1. The two states in the region that did not request a waiver were Nebraska and North Dakota.
2. Information about this study can be found at http://www.ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=333. When the study report is published, it will be available at http://ies.ed.gov/pubsearch/
3. Because of the way the survey was developed and tested, the responses to questions in the importance of feedback characteristics section are not expected to correlate with one another, so summing them would not lead to a meaningful index.
4. Classical test theory examines reliability by focusing on how closely related a set of questions is.
5. Rasch analysis allows for the examination of additional qualities of the information obtained from the survey, including how well questions fit in the survey and how respondents are using the response scale.
6. Confirmatory factor analysis helps examine whether the survey measures the categories that it is intended to measure.
7. Pairwise measures of transition between response categories on the latent scale.

References

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

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2055–2100. http://eric.ed.gov/?id=EJ858247

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

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: The Heritage Foundation. Retrieved June 21, 2012, from http://www.heritage.org/Research/Education/CDA00–09.cfm.

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.

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.

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., & Singer, E. (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

U.S. Department of Education. (2014). *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