NATIONAL COMPREHENSIVE CENTER
FOR TEACHER QUALITY

Linking Teacher Evaluation to Professional Development:

Focusing on Improving Teaching and Learning

MAY 2012

Research & Policy Brief

ACKNOWLEDGMENTS

The authors greatly appreciate the advice of reviewers Joellen Killion, Learning Forward; Tricia Miller, Colorado Department of Education; and Glenda Partee, Center for American Progress. Helpful direction and contributions were also received from Bonnie Jones and Fran Walter at the U.S. Department of Education, and from AIR (and TQ Center) colleagues Matt Clifford, Jane Coggshall, Sabrina Laine, and Angela Minnici.

Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning

This **Research & Policy Brief** was developed to support the efforts of states and districts to inform professional growth decisions and opportunities by strategic use of results from teacher evaluation processes.

May 2012

Laura Goe, Ph.D., *ETS*
Kietha Biggers, *ETS*
Andrew Croft, *ETS*

NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

CONTENTS

INTRODUCTION

Recently, teacher evaluation has become a major focus in educational policy debates and research efforts. This increased attention to teacher evaluation has raised questions about the relationship between evaluation and student outcomes. Rivkin, Hanushek, and Kain (2005) and others have demonstrated with value-added research that there are sizeable differences among teachers in their ability to help students learn at high levels—not just across schools but within schools. These differences lie at the heart of the current approaches to measuring teacher effectiveness through teacher evaluation. The results of these evaluations will be used in many districts and states for accountability purposes to make high-stakes decisions about teachers. Evaluation may be a tool to help teachers improve, but school leaders often lack training in how to use evaluation results to guide teachers toward professional growth. However, both outcomes of evaluation—for accountability and for improvement—rely on the same foundation for making decisions: reliable and valid evidence about teacher performance and student learning.

This paper lays out an informal framework for using evaluation results to target professional growth opportunities for teachers within an aligned system of evaluation, leading to higher levels of teacher practice and student learning. It is based on the belief that evidence collected for teacher accountability can also be used to determine the focus and strategies for professional growth for all teachers, but particularly for teachers who are not meeting expectations in terms of their classroom performance or their students' learning. Building trust and strong relationships among teachers and between teachers and evaluators is critical to ensure that teachers can benefit most from evidence-based conversations, resulting in successful use of evaluation results for teacher learning.

THE USE OF EVIDENCE IN AN ALIGNED TEACHER EVALUATION SYSTEM

The alignment of teacher evaluation results with professional growth opportunities can be considered in terms of the collection and use of evidence. It begins with identifying sources of evidence that will be used for evaluating teachers. Once state and district leaders have determined those sources, they must identify instruments and create processes that will ensure that evidence is gathered with high standards of validity and reliability. Next, it will be necessary to develop appropriate responses to the evidence, such as linking professional development to growth over time, both for individual teachers and for schools and districts. The *Appendix* offers some guiding questions that will help states and districts consider how to use evidence to the greatest effect in an aligned teacher evaluation system.

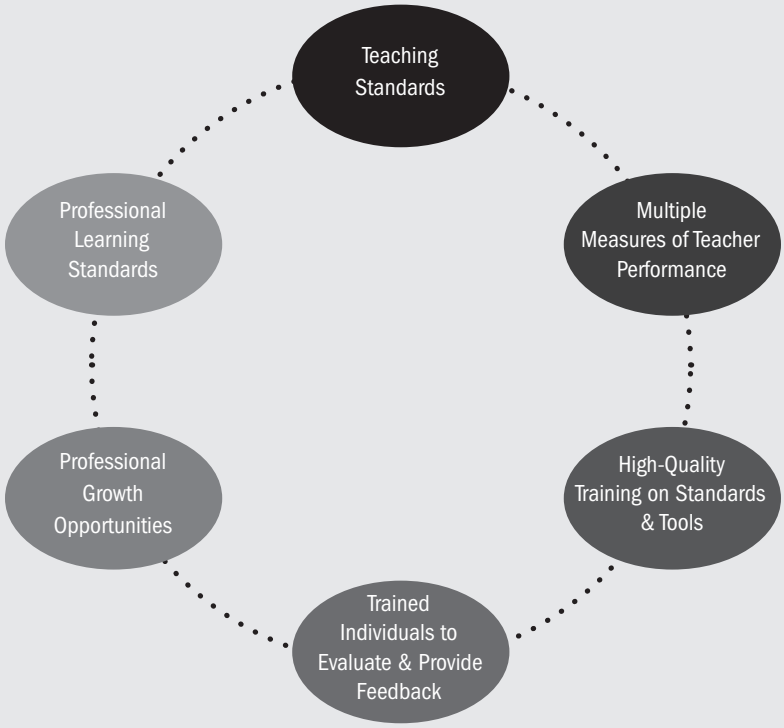
As states develop and implement teacher evaluation systems in response to federal and state priorities, they should consider designing systems that include using evidence gathered through evaluation to inform professional growth. Adding this component to teacher evaluation systems from the very beginning will ensure that implementation decisions support the collection and use of evidence for the purpose of informing professional growth. Described below are six key components to an aligned teacher evaluation/professional growth system.

SIX COMPONENTS IN AN ALIGNED TEACHER EVALUATION/ PROFESSIONAL GROWTH SYSTEM

Evaluation for accountability and for improving performance can be part of the same system. With careful thought and planning in the design stages of a comprehensive evaluation system, education leaders can create an aligned system that meets both goals. The six components that are essential to include in a teacher evaluation system that can be effectively used for professional development are as follows (and shown in Figure 1):

1. High-quality standards for instruction
2. Multiple standards-based measures of teacher effectiveness
3. High-quality training on standards, tools, and measures
4. Trained individuals to interpret results and make professional development recommendations
5. High-quality professional growth opportunities for individuals and groups of teachers
6. High-quality standards for professional learning

Figure 1. Six Components in an Aligned Teacher Evaluation/Professional Development System



1. High-Quality Standards for Instruction

To ensure equity, fairness, and common understanding in teacher evaluation, states need to establish teaching standards, which may also be called standards for instruction, defining criteria for quality teaching. There is no national consensus on what those standards should be; various organizations have proposed standards, and most states have developed their own or adopted national standards, such as the InTASC standards (Council of Chief State School Officers, 2011), or adapted standards from another source such as Charlotte Danielson's Framework for Teaching (The Danielson Group, 2011). Generally, standards should build upon what is known about effective teaching practices, describe what exemplary performance looks like, and serve as a guide for developing the related components. Teaching standards should be compatible with standards for student learning. High-quality teaching standards are a fundamental component of an aligned evaluation/professional development system.

In a well-aligned system, evidence of practice as it relates to high-quality standards will:

- Form the basis for a professional growth plan.
- Give structure and consistency to coaching and mentoring by providing the basis for shared expectations and a common language, and possibly suggesting a direction for development.
- Provide a diagnostic approach to understanding inadequate student learning growth (i.e., determining which standards are not being met and considering how they might relate to student outcomes).
- Offer a set of criteria to help principals, consulting teachers, mentors, and others identify areas in which teachers are successful and areas for improvement.

For example, updated InTASC model core teaching standards focus on common standards of professional practice across grade levels and subject areas for teachers to improve student learning. They include standards in four categories: learning & learners (e.g., learning development), content (e.g., content knowledge), instructional practice (e.g., assessment and planning), and professional responsibility (e.g., leadership and collaboration) (Council of Chief State School Officers, 2011). Notably, these standards include a key feature of professional development: collaboration among teachers. Another noteworthy characteristic is that these standards place equal weight on teacher knowledge of content and learners; they must not only know their subject matter but understand the multiple ways students learn and demonstrate their learning.

When designing and introducing teaching standards, they should be transparent to teachers. They must be clearly defined and communicated to teachers in order for them to know what grounds are being used for their evaluation and to provide direction in their professional learning experiences. Including teachers in the process of adopting standards and designing systems can help ensure that the standards will be meaningful to teachers' daily practice in the classroom.

Ideally, teaching standards should be introduced to teacher candidates in their teacher preparation programs. This introduction helps focus the teacher candidates on key knowledge, skills, and practices that they will be expected to demonstrate in the classroom. Incorporating the teaching standards into the preparation program curriculum will help teachers develop a sense of how the various teaching standards will fit into their own practice when they are in the classroom.

Table 1 below illustrates how teaching standards can be used in an evaluation system to focus on teacher and student success.

Table 1. Using Teaching Standards to Focus on Teacher and Student Success

Actions	Rationale	Application
Teaching standards are created (often by states) and reflect what teachers should know and be able to do.	While some say they "know good teaching when they see it," the standards break down the complex act of teaching into components that reflect specific knowledge and skills that teachers can focus on in order to be successful.	Teachers should be directed to the standards throughout their careers so that they will know explicitly what is expected of them.
Teaching standards are taught to teachers in preparation programs and reviewed with them as part of the evaluation process.	Teacher preparation programs have various philosophies, agendas, and goals, but using teaching standards as a guiding document should be central to all programs. The result will be teachers who enter the classroom knowing what is expected of them and understanding that their achievements as a teacher will be measured by their success on the standards.	Teacher preparation programs should introduce their students to the standards early in order to help them think about their own practice relative to expected and exemplary practice. Evaluators should ensure that teachers know the standards that form this basis for their licensure and evaluation.
Teaching standards are used to guide the creation or adoption of the measures, instruments, and processes of teacher evaluation, coaching, and mentorship.	Once teachers know what is expected of them, they need to know what to do to demonstrate success. The measures, including the evidence that they are designed to collect, are crucial to helping teachers see the link between their classroom performance and the evaluation results. Rather than evaluators making an arbitrary, subjective judgment on their performance, teachers can see that evaluation results are based on evidence linked explicitly to the standards.	Measures should be linked to the specific standards on which they are designed to collect information. If what is valued is accurately reflected in the teaching standards, then teacher performance on standards should accurately identify effective teaching and less successful teaching.
Evidence of performance on teaching standards is collected as part of the teacher evaluation process.	Different measures may be useful for collecting evidence on different teaching standards. When teachers understand the relationship among standards, evidence, and evaluation results, they can become active and engaged participants in the process of collecting and providing high-quality evidence to ensure accurate evaluation results.	Each measure used should specify how the evidence collected reflects teacher performance on a specific standard.
Evidence of teacher performance on standards is used by evaluators to discuss teacher practice and areas where growth is needed.	Evaluators should assume that teachers are doing the best they know how to do. If the best is not good enough, then part of the evaluator's job is to help provide teachers with information that will enable them to improve their practice to bring them closer to successfully meeting all of the standards. Evidence is at the heart of this conversation.	Through a discussion of the evidence with the evaluator, teachers can come to a better understanding of their strengths and weaknesses so that they can focus on addressing them.

Actions	Rationale	Application
Evidence of student performance on subject and grade standards is included in discussions between the evaluator and teacher as it is a reflection of successful (or less successful) practice.	The national policy changes and subsequent culture shift taking place in schools has happened quickly and taken many educators by surprise. No longer are teachers being judged solely on their practice; now, their students' learning growth is included in evaluation scores. However, teachers want to succeed, and by making them partners in the evaluation process, student assessment can play a role in teaching success.	Teachers should consider evidence of student learning growth in order to impact student results in positive ways. Regular assessment of student progress against learning standards allows the teacher to have immediate impact on learning outcomes.
A growth plan based on teacher performance on standards is developed collaboratively between the teacher and his or her instructional manager, principal, mentor, support provider, coach, or consulting teacher.	For teachers to improve their practice, they need to know what they are doing well, and what they need to work on. A standards-based professional growth plan focuses teachers on the specific standards where they are not meeting expectations.	Once the evidence on teacher performance has been discussed in detail, the teacher and his or her support provider create a plan that will allow the teacher to consider what the evidence means in terms of specific knowledge or skills that the teacher needs or must improve upon. Some teachers may require comprehensive growth plans that focus on improvement across many standards. For teachers who are already performing well on most standards, a growth plan may be focused on a single area where they wish to further their development.
Subsequent teacher evaluations should revisit previous standards-based evidence in order to judge growth in practice.	Teaching is by nature cyclical: Teachers offer instruction and provide activities to help students learn, evaluate the results of that process, and make adjustments or adaptations to improve results. Teacher evaluation should be similarly cyclical: Teachers are evaluated, evaluators provide feedback and recommendations, and teachers apply the recommendations to their practice. Evaluation allows teachers to demonstrate their success. Ongoing feedback allows teachers to make additional adjustments and demonstrate even greater success in subsequent evaluations.	Evaluators should identify a mechanism for keeping evaluation results in a single longitudinal database. This will make it possible to judge teachers' growth over time, particularly on standards that have previously been identified as areas for improvement.

2. Multiple Standards-Based Measures of Teacher Effectiveness

During the design phase of creating an evaluation system, developers should think carefully about the types of evidence that will need to be collected to measure teachers' performance against the teaching standards. When they agree on the types of evidence needed, then they can make decisions about what measures to use to collect that data. Measures used for teacher evaluation are typically designed to result in scores that assign teachers to a level along a performance continuum, such as "beginning" or "accomplished" teaching. For a discussion or how single scores can be created from multiple sources of evidence in evaluation systems, see Leo and Lachlan-Haché (2012).

Using multiple standards-based measures affords the opportunity to triangulate results (i.e., compare scores across different types of measures to see if all the scores consistently support a single performance level for that teacher). Triangulating results allows for greater confidence in any one score when all scores are not in agreement. It also makes it possible to identify problems in the collection or

interpretation of evidence when the scores are not in agreement. Scores for an individual teacher that are very high on a measure of student learning growth but very low on a classroom observation measure should raise a red flag that there is probably a scoring or interpretation error as good learning results are assumed to be the result of good teaching.

There is wide variation in what teachers teach and who is being taught. When choosing measures of teacher quality, It is important to realize that some measures are more useful than others when it comes to helping teachers improve their practice, as shown on pages 10–19. In an analysis of 34 professional development studies and 11 reviews of professional development, van Veen, Zwart, and Meirink (2011) concluded that effective professional development "…should be related to classroom practice, more specifically to subject content, pedagogical content knowledge, and student learning processes of a specific subject. When teachers develop with respect to these aspects of content, an increase in teacher quality and student learning results." (p. 17)

Another important feature of measures is their sensitivity to student populations as it relates to teaching. For example, different rubrics may help collect specific types of evidence for teachers of English language learners or students with disabilities (Holdheide, Goe, Croft, & Reschly, 2010). Similarly, the validity of measures may be impacted by their appropriateness for the context in which they are used. Later in this section, five of the most common teacher evaluation instruments will be discussed, but first, five general criteria are offered to assist developers when making decisions about which measures to include in their evaluation systems:

- Measures are directly and explicitly aligned with teaching standards. This alignment ensures that what is valued most is being measured and what is expected is unambiguous.
- Measures include protocols and processes that teachers can examine and comprehend. Evaluation that makes sense to teachers will be more meaningful and have a greater impact.
- Measures allow teachers to participate in or coconstruct the evaluation. Collecting evidence on themselves encourages reflection on practice and empowers teachers to be proactive in their evaluation.
- Measures allow teachers opportunities to discuss the results with evaluators, administrators, colleagues, teacher learning communities, mentors, and coaches. Active intellectual engagement leads to deeper learning.
- Measures align with professional development offerings. The type of data collected lends itself to informed professional development decisions.

> Multiple measures paint a more complete and elaborate picture of a teacher's strengths and weaknesses, ensuring better alignment with professional growth opportunities.

We will keep these principles in mind as we describe to what degree the following common measures can contribute to an aligned system that supports improved teaching and learning:

- Classroom observations
- Student learning growth
- Portfolios (traditional or digital)
- Student surveys
- Classroom artifacts (work samples)

No one measure can provide all of the information needed to accurately assess a teacher's performance for accountability or professional learning purposes. Using multiple measures of teacher performance will be important to understanding the full range of a teacher's abilities (Kane & Staiger, 2012). Multiple measures paint a more complete and elaborate picture of a teacher's strengths and

weaknesses, ensuring better alignment with professional growth opportunities. The evidence-collection tools and scoring rubrics associated with different measures then serve to define expectations, justify scores, and create the opportunity to "diagnose" and target areas where professional growth is desired. It is important to note that understanding a teacher's strengths is valuable as well when assessing in-school expertise to support professional development planning.

Classroom Observations

Classroom observations can be a valuable part of a performance evaluation as well as one of the best sources of information to guide teacher professional growth. In particular, observations provide useful data on teacher-student relationships and the learning environment. Classroom observations alone are unlikely to lead to improved teaching and learning, however, without discussions about the standards assessed and evidence obtained in the observations. In addition, including evidence of student learning in these conversations can lead to teachers focusing primarily on whether students are actually learning rather than focusing only on their own instruction.

While teacher-student relationships have been seen as key to student success, there are a number of studies that emphasize the importance of that relationship. A recent study on teacher-student relationships confirmed that positive relationships could influence academic engagement and performance, particularly for at-risk students (Roorda, Koomen, Spilt, & Oort, 2011). Most classroom observation instruments, such as Charlotte Danielson's Framework for Teaching (2007), focus at least in part on teacher-student relationships, while the CLASS™ [Classroom Assessment Scoring System] instrument (La Paro, Pianta, & Stuhlman, 2004) makes them the major focus of the observation.

The learning environment created and maintained by the teacher and the interactions between the teacher and the learners are essential elements for student learning to take place. These elements can best be assessed through direct observation. Most other measures in teacher evaluation are indirect: the artifacts of instruction (such as lesson plans and student work samples), student reports (via surveys), teacher self-reports (such as portfolio entries), and the outcomes of instruction (student learning growth).

With information from observations, teachers and observers can focus on evidence of both positive and negative interactions with students and develop strategies to expand the positive interactions and reduce the negative ones. Engaging in "evidence-based" conversations with observers encourages teachers to be reflective about their practice. Conversations between observers and teachers have long been used in the "lesson study" model of teacher professional development (Rock & Wilson, 2005) as well as in formal or informal peer observations (Joyce & Showers, 1988).

Charlotte Danielson's Framework for Teaching (Danielson, 2007) is an observation instrument that provides a foundation and shared language for having professional conversations, but Danielson also encourages that teachers be given the notes of the observation and time to reflect on their own performance prior to the conversation. This preparation allows teachers to be fully engaged in their own evaluation and active contributors to the discussion on professional learning needs.

Evidence of student learning coupled with observation data can inform decisions about professional learning. However, this link is only as strong as the quality of observer training. Results from research on how well Framework for Teaching scores correlate with student growth (measured by standardized test scores) has been weak until recently, which may reflect poor evaluator training. When well-trained and calibrated observers are used, the correlations

are decidedly stronger. Kane, Taylor, Tyler, & Wooten (2010, 2011), using rigorously collected classroom evaluation data from Cincinnati schools, found that classroom practice evaluation scores could reliably predict differences in student achievement growth in reading and math. Further, while teachers' overall scores were most strongly correlated with student achievement growth, individual classroom practices, such as classroom management and questioning, were also correlated.

Another recent study by the Consortium on Chicago School Research at the University of Chicago found that "…classroom observation ratings were valid measures of teaching practice; that is, students showed the greatest growth in test scores in the classrooms where teachers received the highest ratings on the Danielson Framework" (Sartain, Stoelinga, & Brown, 2011, p. 2). Results from the Gates Measures of Effective Teaching (MET) study (Bill & Melinda Gates Foundation, 2010) have also provided evidence of correlations between student achievement gains and all five of the observation measures used in the study. Results from these recent studies are important because they suggest that there are specific teacher practices and areas of knowledge that are observable and predictive of student achievement growth. By leveraging results of classroom observations to give teachers specific feedback on these components and align growth opportunities with results, teaching and learning should improve. Commonalities of observations and other standards-based measures of teacher effectiveness support a well-aligned system of teacher evaluation and professional learning.

The type of feedback provided to teachers and how it is given often varies depending on whether its purpose is summative (evaluation) or formative (professional growth). Coaches, teacher leaders, and others can use results shared with teachers as part of summative feedback as a springboard for formative conversations. For that reason, evaluation documents from observations, review of portfolios, or reports on student growth should include specific written feedback focused on the scores teachers received and the basis (evidence) for those scores. Formative feedback is appropriately used by teacher leaders, principals, and peers to provide ongoing guidance for improving practice and student outcomes. Formative feedback should include multiple opportunities for discussion throughout the school year (rather than at the end of the year as with summative feedback) and may be focused on the teacher's priorities for growth as well as the results of evaluation cycles.

Student Learning Growth

Student growth data can be used to inform instructional decisions and guide teacher practice. In a committee report commissioned by the Board on Testing and Assessment of the National Research Council, Pellegrino, Chudowsky, and Glaser (2001) contend that "(1) something important should be learned from every assessment situation, and (2) the information gained should ultimately help improve learning" (p. 7). Fullan, Hill, and Crévola (2006) state that the key to transformation of teacher practice "…lies in the smart use of data to drive instruction" (p. xvi). Fullan et al. also note that while many school systems collect and even analyze student growth data for the benefit of schools, teachers do not know how to use the data to improve instruction. They sum up the problem as follows: "Even if the data are better analyzed, teachers do not know how to translate the information into powerful, focused instruction in response to individual students' needs" (Fullan et al., 2006, p. xvi). Thus, even if we have "actionable" data collected for various purposes, including evaluating teachers, there is still a need for guidance in how to use that data. Unfortunately, student growth data is of little use in improving instruction without specific guidance teachers can use to change their practice and develop new teaching strategies.

The use of student learning data to improve instruction is strongly recommended in a review of research evidence. Hamilton et al. (2009) described a number of reasons why data should be used in instructional decisions, including "gauging the instructional effectiveness of classroom lessons" and "refining instructional methods" (p. 5). With data on individual students' performance, teachers and those who support them can consider these results in light of other evidence, such as results from teacher observations, to determine whether there are specific strategies or approaches that might improve student learning.

Only high-quality student growth data should be used for purposes of informing teacher professional growth. According to Peine (2008):

> If student data represent multiple sources of information, if they present reliable trend data, if they produce accurate inferences about student achievement and program concerns, then they can and should play an important role in targeting professional growth. If, however, the quality of the data and/or their interpretation is lacking or poor, then their use in identifying areas of need of professional growth should be limited. (p. 54)

Furthermore, student growth data varies in how well suited it is for use as a tool for teachers' professional growth. In order to be most useful, evidence of student learning should be directly connected to specific content as well as to the delivery of that content (i.e., instructional practices). When examining evidence of student learning, teachers and those who support them should, with proper training, be able to identify trends in overall learning, gaps where content is not being mastered, and areas where students are demonstrating proficiency. Through connecting this information to both the content and how the content was taught, it should be possible to make some determinations about which strategies were most and least effective. In some cases, this process will require considerable analysis and effort to make reasoned determinations about what works best, but there are some programs that have been designed to make this link between teaching and learning more explicit and tie it to specific professional growth opportunities. One example is the STEP™ (Strategic Teaching and Evaluation of Progress) Literacy Assessment designed by the University of Chicago Urban Education Institute and implemented in early grades in Chicago Public Schools. STEP uses targeted literacy assessments in combination with a data management system and professional development (see sidebar for more information).

STEP LITERACY ASSESSMENT TOOL

The STEP Literacy Assessment™ is a developmental literacy assessment, instructional tool, and data management system that defines the pathway and tracks the progress of prekindergarten through third-grade students as they learn to read using research-based milestones.

STEP™ enables educators to implement a developmental approach to teaching reading, using evidence to inform instruction and introducing targeted interventions based on that evidence.

The University of Chicago Urban Education Institute (UEI) has developed STEP™ over the last decade and has worked with Chicago Public Schools and others in studying its impact. STEP™ is most appropriate for districts and schools that are practicing small-group differentiated instruction and are interested in the use of formative assessment data to inform instructional improvement.

The STEP™ program has three elements:

- The **STEP™ Assessment Kit**: A toolbox that provides teachers with the materials necessary to administer the assessment.
- The **STEP™ Tool**: A powerful data management system for teachers, principals, and districts to visualize student, teacher, and school performance.
- **STEP™ Guidance**: Ongoing professional development sessions that support and guide users in optimal classroom literacy instruction.

Originally developed to support literacy learning in a small number of public schools, STEP™ has evolved into a powerful assessment tool and performance management strategy that is being used in schools and districts across the country.

Portfolios

Portfolios can provide teachers with an opportunity to be full participants in the evaluation process. Portfolios require teachers to reflect on various aspects of their practice, and that reflection can be a powerful source of teacher learning. However, their usefulness for teacher evaluation and for determining professional growth needs depends on the guidance provided to teachers on what should be included in them and why. For example, with little or no guidance, teachers tend to include evidence that they anticipate will present a favorable impression to evaluators such as compliments from parents, examples of stellar student work, or evidence of participation in a conference. For a portfolio to be useful in planning professional growth activities, it needs to be focused on instructional data about student learning, instructional challenges and how they were addressed, and reflections on practice. These types of entries enable principals, coaches, mentors, and peers to have meaningful conversations about teachers' specific needs and how they might be addressed through professional growth opportunities. A portfolio stuffed with commendations generates little in the way of constructive discussion that could improve practice.

Student Surveys

Student ratings of teachers have been found to correlate with measures of student achievement (Bill & Melinda Gates Foundation, 2010; Wilkerson, Manatt, Rogers, & Maughan, 2000) and can serve as a useful source of feedback on teacher practice (Peterson, Wahlquist, & Bone, 2000). One caveat in using student surveys for evaluation purposes is that they may require multiple years of data collection to yield stable results (Peterson et al., 2000). Another caution is that research on using student surveys for teacher evaluation is in its infancy. Furthermore, surveys differ greatly in focus and quality and may also vary in how useful they are to teachers in terms of improving their practice.

From a practical standpoint, surveys are relatively inexpensive to conduct and analyze and they provide low-cost information that can be triangulated with other measures of teacher practice and student outcomes as part of a multiple-measures evaluation system. There is simply no other measure that can capture multiple students' viewpoints as efficiently as a survey. While surveys are worth considering as one measure in a system of multiple measures, they have an added benefit of providing information about teacher practices, such as emotional and instructional support provided to students, which can be helpful to teachers and to those who support their professional growth. Research on surveys has shown that students can distinguish between emotional support and instructional support provided by teachers. While younger students value emotional support more, older students value instructional support more, though these two aspects of teaching are not mutually exclusive for either group (Peterson et al., 2000). However, the usefulness of student surveys for improving teaching practice depends in part on teachers' openness to students' feedback. Some teachers may feel threatened by the survey, particularly if they feel that students will not evaluate them fairly. These fears may be alleviated by giving teachers an opportunity to examine the survey questions, which are designed to collect specific information about the learning environment and instructional practices rather than students' opinions of their teachers.

Using student surveys to guide decisions about professional development can be done in several ways. As a whole-school, team, subject, or grade activity, aggregated results can be examined across classrooms to get a general sense of students' perceptions about their classroom experiences. In fact, Harvard's Tripod Student Survey (see sidebar) was designed with this whole-school approach—using aggregated data for purposes of discussion. Examining aggregated data is less threatening to teachers because the discussion centers on data from multiple sources, and problems and solutions are

considered for all participating teachers as a group. Results can also be examined one on one with a consulting teacher, instructional coach, or other professional who can help the teacher focus on specific aspects of the results and develop a plan of action to address them. Finally, results can be examined as part of a professional learning community. In this activity, teachers would use survey results as the beginning of a concerted effort to improve the learning environment, time on task, classroom management, or other aspects of instruction and practice that are in need of improvement based on survey results. Focusing on teaching standards and student learning standards, the teachers would collaborate on a plan to address the problem area(s), implement the strategies, and conduct the survey again to measure improvement. Reflecting on the process and the learning that has taken place, and describing next steps would complete the cycle.

TRIPOD STUDENT SURVEY

One example of a student survey that is gaining attention is the Tripod Student Survey Assessment developed by Harvard researcher Ronald Ferguson and administered by Cambridge Education. Part of the larger Measures of Effective Teaching (MET) study, financed by the Bill and Melinda Gates Foundation, examination of the Tripod survey shows that student perceptions of their instructional environment can be a useful measure of teacher quality in a multiple-measures system (Bill & Melinda Gates Foundation, 2010). In six districts across the U.S., researchers for the study collected data regarding the practices of mathematics and English language arts teachers in grades 4–8 to look for relationships among different measures of teacher quality. One compelling early finding is a predictive relationship between teacher scores on the Tripod survey and student performance on state standardized tests. In other words, a teacher's survey results in one class were related to the achievement gains of his or her students in another class. Further, there is agreement between groups of students on the strengths and weaknesses of a given teacher, suggesting that the survey is also a reliable measure of student perception. Correlations were seen most strongly in the areas of classroom management and instructional rigor (Bill & Melinda Gates Foundation, 2010). (More information on the MET study can be found at http://www.metproject.org/.)

The Tripod survey uses research-based indicators to assess students' attitudes, experiences, and perceptions in the classroom as they relate to teaching and climate. Questions are focused on specific behaviors of students and teachers to avoid overly subjective responses that tell more about whether or not a student likes a teacher rather than a teacher's instructional competence. Questions are organized into three domains of teacher effectiveness—content knowledge, pedagogic skill, and relationship-building skills. These are then further divided into seven multiitem measures referred to as the "7 Cs." These are:

- Caring about students (nurturing productive relationships)
- Controlling behavior (promoting cooperation and peer support)
- Clarifying ideas and lessons (making success seem feasible)
- Challenging students to work hard and think hard (pressing for effort and rigor)
- Captivating students (making learning interesting and relevant)
- Conferring (eliciting students' feedback and respecting their ideas)
- Consolidating (connecting and integrating ideas to support learning)

(For more information about the Tripod survey assessments, visit http://www.tripodproject.org/uploads/file/assessments-flyer.pdf.)

Surveys are designed for three levels, grades K–2, 3–5, and 6–12. They also provide analysis and reporting to client schools, and strategic planning and professional development tied to understanding and addressing the results.

Students spend more time in the classroom than any other education stakeholder and are perhaps in the best position to offer insight into a teacher's skills. To ensure that their feedback is most useful, it is important that the survey be administered within a culture of trust and openness. Teachers must be open to their students' feedback and trust in their sincerity, and students must trust that they will be taken seriously and that their confidentiality will be maintained. Under these conditions, student surveys offer valuable data that could be used to inform teacher professional development. However, like all evidence of teacher performance, teachers should have an opportunity to see how they will be evaluated, reflect on their results, and make decisions about their professional learning needs.

Classroom Artifacts

Classroom artifacts typically include teacher assignments and resulting student work. Teacher work samples document teachers' ability to teach according to standards and impact student learning (Schalock, 1998). A study by Denner, Salzman, and Bangert (2001) looked at whether work samples could be used to differentiate among teachers and found evidence to suggest that they could be used for that purpose. As a measure suitable not only for assessing teachers but for providing growth opportunities, artifacts are very promising. Artifacts have the advantage of providing considerable detail about students' understanding of particular concepts. Discussing artifacts with others in a lesson study format or teacher learning community, or meeting one on one with a consulting teacher, curriculum specialist, assessment specialist, or instructional coach, a teacher can ferret out misconceptions or understand more about the thought processes that led to a particular answer, whether right or wrong. This information can then be used to guide teacher practice and instructional decisions in ways that should lead to improved student learning outcomes. How a teacher plans for and manages this task is a source of evidence for evaluation represented by the student artifact.

The collection of teacher assignments and resulting student work could be a measure used within a comprehensive teacher evaluation system strictly for evaluating teacher performance. However, examining teacher assignments and resulting student work could also be done as part of ongoing professional development. Matsumara and Pascal (2003) specifically support the idea of using "collaborative professional development" around classroom assignments, providing opportunities for teachers to consider their assignments in terms of both content and implementation of assignments, and focusing on students' opportunities to respond to those assignments with high-quality work.

However, for teachers to realize the full value of professional learning around assignments and artifacts, analysis must focus on students' progress towards appropriate grade and subject standards.

3. High-Quality Training on Standards, Tools, and Measures

Effective training is essential to ensure that observers or evaluators are familiar with the standards being measured, the evidence to be examined, and how to appropriately score the evidence. The term "training," as it is used throughout this brief, does not indicate a one-time occurrence, but a continual refinement of understanding how standards, measures, and tools can contribute to an alignment between teacher evaluation and professional development. Through training, stakeholders gain a better understanding of the purpose and expectations embedded in the system. If component 1 (high-quality standards for instruction) and component 2 (multiple standards-based measures of teacher effectiveness) are in place, then this understanding leads to greater buy-in among teachers and administrators, which is important for supporting professional development. Researchers can point to at least five characteristics of high-quality professional development for teachers (discussed within component 5), but what underscores all of them is teacher buy-in (Archibald, Coggshall, Croft, & Goe, 2011). This objective should be a focus from the beginning of designing and implementing teacher evaluation systems that improve teacher practice. Ideally, teachers will be involved at every level, providing their insights and expertise, but at a minimum, teachers must receive high-quality training on standards and measures to begin promoting investment in and ownership of their professional development. For any system to be successful, its primary stakeholders—in this case teachers—must believe in its value.

Without standards and standards-based measures, there is no evidence, only opinion. And without training, scores on observations and other measures are based on personal judgment rather than evidence. The result is that the same teacher could get very different scores depending on who observed him or her. To ensure the validity of the evaluation system, evaluators and teachers should be trained on both the standards and the measures. Both teachers and evaluators need to understand and agree on what the standards mean and look like in practice. Training also increases the likelihood that scores from different evaluators will be similar (it improves the reliability of the instruments). A strong evaluator training and certification process not only protects the system from fundamental flaws, it also serves as excellent professional development, particularly for novice teachers and those new to the evaluation system. Including peer evaluators in the system, an approach that has been successfully used in many districts and schools, also opens up possibilities for teacher growth and development. Anyone has the potential to be an evaluator—principals, teachers, instructional leaders, and coaches—as long as they are provided with sufficient training, pass a certification test if required, and periodically calibrate their skills with other evaluators to maintain the quality and reliability of the system.

4. Training to Interpret Results and Make Professional Development Recommendations

This is the critical role that principals typically play, yet are often not adequately prepared to undertake. Administrators and teachers do not always receive sufficient training in using results from evaluation and student outcomes to identify strengths and weaknesses in instructional practices, the learning environment, or classroom management. If they were not trained in their preparation or leadership programs, there may be a need for such training so that they can learn to use evidence and results to guide decisions about areas where teacher growth is needed. Other support personnel may also need such training, such as evaluators, instructional leaders, coaches, and/or other support providers. Ideally, it is the evaluators who are able to provide this guidance at the time of feedback. In any case, the purpose of this role is analyzing the evaluation results to make decisions about professional learning needs at the teacher, school, and district levels.

Often, professional development decisions are guided by school or district priorities. For example, a district that has a goal of improving reading may emphasize professional development focused on reading across the curriculum and similar initiatives to ensure that all teachers are prepared to teach reading in every content area. As evaluation results are interpreted, they can be specifically tied to district and school initiatives and goals. Creating a classroom learning environment conducive to high levels of success in reading would tie in well with the reading goal in the example above. Various types of evaluation instruments, such as classroom observations and student surveys, can be used to examine how well teachers have created and maintained such learning environments. Tools such as iObservation (Learning Sciences International, 2011) directly link observer ratings with professional development resources, such as books or curricular materials. It is even possible for results to be directly linked to exemplars of teaching practice recorded in video. The Classroom Assessment Scoring System (CLASS) is another instrument that ties results to specific professional development recommendations. However, it must be emphasized that evaluators or instructional coaches would need to discuss results and recommended professional learning with the teacher to ensure the greatest benefits to teaching and learning.

Role of Leadership

School leaders (e.g., principals, administrators, teacher leaders) have many roles within the school, including instructional leadership. In many states and districts, principals are evaluated on many factors, including the success of their teachers and students. Thus, they must have or develop the ability to serve as instructional leaders and guide teachers towards improved practice that will ultimately lead to better student outcomes.

Given the many roles that school leaders play—building supervisor, personnel manager, budget director, disciplinarian—it is challenging for them to devote the significant time needed for comprehensive teacher evaluation. However, the role of instructional leader comes with certain requirements, including gaining a thorough knowledge of the professional growth needs of the teachers in the building. Moreover, principals have to be proficient at identifying opportunities for teachers to meet those professional growth needs—in school-based communities of practice, with mentors/coaches, or with external supports. Given time pressures, it may be tempting for school leaders to rely on technology to improve the efficiency of the evaluation process. However, technology is no substitute for providing leadership in meeting the professional growth needs of teachers. Instead, school leaders should consider distributive leadership, sharing teacher evaluation with other instructional leaders in the building, including assistant principals and lead teachers.

With the rapidly changing mandates around teacher evaluation, it is highly likely that school leaders will need training around such areas as instructional standards, measures used in evaluating their teachers, and professional growth opportunities. They may need guidance in how to have productive discussions with teachers about student assessment results or how to create a learning environment that supports student success. It is essential that principals ensure that evaluation of performance and resulting feedback are given in a climate of mutual respect and trust, which may require specific training. While they will likely be evaluated on school outcomes such as student achievement, school leaders are also likely to be evaluated on how they support teacher professional growth and work collaboratively with teachers to improve student learning.

Importance of Feedback

Little (2006) reported results from a longitudinal case study of schools and noted the importance of "…focused and timely feedback on individual performance and on aspects of classroom or school practice" (p. 22), an element that was notably present in successful schools, some of which served at-risk students. It is through discussing and reflecting on evaluation results that evaluation becomes a system that supports professional learning, not just accountability. Actively engaging teachers through self-reflection and professional discussion makes evaluation more constructive for them and supports what we know about learning— that the learner must do most of the work (Danielson, 2010).

Charlotte Danielson, author of *Enhancing Professional Practice: A Framework for Teaching* (2007) and education consultant, emphasizes the importance of the feedback conversation to professional learning. Many schools and districts have used the Framework for Teaching as a basis for their teaching standards and evaluation, but as she notes on her website: "The Framework may be used for many purposes, but its full value is realized as the foundation for professional conversations among practitioners as they seek to enhance their skill in the complex task of teaching." (For more information, visit **http://charlottedanielson.com/theframeteach.htm**.)

When feedback discussions are included, the evaluation process itself has been shown to have a sustained impact on teacher practice prior to any prescribed professional learning.

Taylor and Tyler (2011) found when studying Cincinnati Public Schools' long-running Teacher Evaluation System (TES) that "high-quality classroom observation-based evaluation improves mid-career teacher performance, both during the period of evaluation and in subsequent years…" The system studied includes a feedback session conducted by a trained evaluator; this opportunity to receive and discuss expert feedback may well be the key factor in improved practice. The feedback provided should be intentionally focused on improvement and offer recommendations for change in practice (Danielson, 2010; Milanowski, 2004). The resources (personnel time and other costs) used for the evaluation process are well spent as the evaluation can provide important evidence that can be used to better align professional growth opportunities. For example, if evaluators observe a pattern, such as new teachers having difficulty with aspects of classroom management, that information can be used to trigger coaches, mentors, and principals to provide specific feedback related to classroom management. The process of evaluation provides the evidence of areas where teachers need help, but that process alone does not change teaching practice. Rather, using the evidence for professional growth opportunities and coaching sessions is where it will have an impact on instruction and student outcomes. Observation alone without the opportunity for feedback and discussion may serve accountability purposes but will have little or no impact on teaching and learning.

Even when the measures used in an evaluation system are effective, sharing results from those measures with teachers is a skill that is not well developed in many evaluators. Sartain et al. (2011) found that many principals lacked appropriate coaching skills to help teachers improve instructionally. A key component of a successful feedback system should include training for those who will be conducting conferences with teachers to discuss evidence collected during the evaluation process.

As a supplement to principal/evaluator feedback, technology may play a role in providing teachers with specific feedback on their lessons. My Teaching Partner (see sidebar) is an example of a web-mediated coaching program in which teacher consultants provide feedback on videos of lessons. This approach may be particularly helpful in schools where additional coaching and support are needed or for teachers who need assistance with specific areas of their practice.

STUDY OF MY TEACHING PARTNER–SECONDARY

When evaluation is aligned with professional development opportunities, we begin to see its formative uses and how evaluation itself becomes a form of professional development. A study by Allen, Pianta, Gregory, Mikami, and Lun (2011) illustrates this point. They examined the efficacy of a teacher professional development program, My Teaching Partner-Secondary (MTP-S), which is based on the evaluation tool CLASS-S (Classroom Assessment Scoring System-Secondary). MTP-S is a web-mediated coaching program focused on enhancing teacher-student interactions to increase student motivation and engagement. Trained teacher consultants evaluate teachers by watching videos of lessons and using the CLASS-S domains to determine strengths and weaknesses in student-teacher interactions. Consultants provide feedback via a secure website, and teachers are encouraged to reflect on the lesson. This is followed by a 20- to 30-minute phone conference between teacher and consultant. The process repeats twice each month throughout the school year.

Using an experimental design and working with 78 secondary school teachers, Allen et al. were able to show that MTP-S had a positive impact on student achievement. Teachers in the treatment group saw an average increase in scores from the 50th to 59th percentile. It is interesting to note that this gain was seen in year two of the study after teachers had completed the program in year one. Authors attribute this finding to the accumulated effect of the professional development, supporting studies that suggest, to be successful, professional development must be sustained over time and include continuous feedback (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007).

5. High-Quality Professional Growth Opportunities for Individuals and Groups of Teachers

The fifth component of an aligned evaluation/professional development system is ensuring that teachers have access to high-quality professional development. There is limited research on what characteristics professional development should have in order to impact teacher practice and instruction in ways that translate to improved student learning.

An important step in planning for professional growth is to take inventory of current professional development opportunities and make decisions about what should be continued, what should be eliminated, and where gaps exist. There are two main points to keep in mind when making these decisions: the characteristics of high-quality professional development and the meaning of the scores that will be produced by the selected evaluation tools.

Recent research has suggested particular features of professional development that are most effective. Archibald et al. (2011) summarize these features as follows:

- Alignment with school goals, state and district standards and assessments, and other professional learning activities including formative teacher evaluation
- Focus on core content and modeling of teaching strategies for the content
- Inclusion of opportunities for active learning of new teaching strategies
- Provisions of opportunities for collaboration among teachers
- Inclusion of embedded follow-up and continuous feedback

It should be emphasized that job-embedded professional development (JEPD) holds considerable promise for improving instruction and student outcomes. JEPD can also play an important role in providing the structure and continuity for teachers to use results from classroom observations and other measures to make changes in their day-to-day practice that will impact student learning (Croft, Coggshall, Dolan, Powers, & Killion, 2010).

Identifying and providing professional learning opportunities that exemplify the above characteristics is important to any educator performance management system. However, for evaluation to be aligned with professional development, an important feature of a district's professional growth plan is that it addresses identified needs in teacher practice and/or schoolwide goals for improvement. It is through valid and reliable evaluation that these needs are identified. While it would be hard to determine exactly where the needs lie while planning for a comprehensive evaluation system, with a deep knowledge of standards and evaluation protocols, developers can ensure access to relevant professional development when evaluation scores suggest a particular need, both individually and schoolwide.

However, before those connections can be made, district leaders must lay some groundwork. Specifically, before an aligned evaluation/professional development system can be implemented, districts should have already mapped out a professional learning program and the mechanisms for accessing it. This process includes:

- Identifying research-based best practices that fit the local context, such as establishing teacher learning communities that focus on content and pedagogy.
- Tapping in-district expertise to provide professional development focused on high-quality instruction, skilled classroom management, and the use of assessment results to differentiate instruction. Memphis City Schools have been using this approach for a number of years (see sidebar on next page).

- Finding outside supports to fill gaps in internal capacity; for example, identifying external experts who could run a group professional development session on an evidence-based instructional strategy for math or a local college that offers high-quality professional development that aligns with school goals.
- Creating a mechanism that will connect teachers to relevant professional learning opportunities based on evaluation results.
- Preparing for and implementing structural changes in the school week/calendar to accommodate increased teacher participation in professional learning, such as weekly time set aside for professional learning communities.

This process may mean simply reviewing a district's current professional development program, aligning supports already in place with standards, and filling any gaps with in-house and external resources, or it may mean a complete overhaul of the professional development system. Structural changes in the central office to support the alignment of evaluation results and professional development may be needed, such as ensuring that those in charge of professional development for the district have access to aggregated teacher evaluation results in order to prioritize and plan districtwide professional development offerings. A district-provided "dashboard" to house and access evaluation results, as well as aligned professional growth plans and perhaps indicators for teachers' progress in professional learning, may be a way for districts to support these efforts.

Aligning and describing the various aspects of a professional learning program during the design stages of a comprehensive evaluation system will go a long way to support smooth implementation. However, to ensure that the program is able to do what it is intended to do—improve teacher practice—it is imperative that school leaders pay attention to building trust and strong relationships among teachers and between teachers and evaluators, so that potentially difficult conversations can be conducted in both nonthreatening and productive ways. For more information on how this can be done, see *Everyone at the Table* (**http://www.everyoneatthetable.org/**). This online resource provides materials to genuinely engage teachers and stimulate candid and respectful dialogue that gets to the heart of their perspectives on teacher evaluation.

MEMPHIS TEACHING AND LEARNING ACADEMY

The Memphis Teaching and Learning Academy is a professional development delivery system for Memphis teachers. Founded in 1996, it is nationally commended and particularly notable for the careful alignment of teaching standards, evaluation, and professional development. The observation instrument is aligned with the state teaching standards, and professional development offerings are aligned with specific indicators from the evaluation instrument. When teachers receive scores from an observation, they can quickly find specific professional development offerings that are linked to specific indicators. If there is an indicator on which their performance was less than satisfactory, they can focus on pursuing the professional development offerings aligned with that particular indicator. If they are already rated excellent but there is an area that they have a desire to grow in (such as using technology in the classroom), they can elect the appropriate professional development options.

It is also notable that most of the professional development offerings are created and delivered by Memphis practitioners and teachers who propose and develop courses. This allows Memphis City Schools to take advantage of the local leadership and expertise, and reduces professional development costs. Moreover, many of the offerings are videotaped and made available online so that teachers can access them at their convenience. In a recent professional development catalog, there were 336 courses, many offered online. For more information, see the following:

- *Main Site:* **http://www.mcsk12.net/admin/tlapages/academyhome.asp**
- *Professional Development Catalog:* **http://www.mcsk12.net/aoti/pd/docs/PD%20Catalog%20Spring%202011lr.pdf**
- *Individualized Growth Resource Book:* **http://www.mcsk12.net/aoti/pd/docs/Individualized%20Growth%20Resource%20Book.pdf**

Uses of Technology

Utilizing technology when collecting and analyzing evidence of teacher effectiveness can be helpful in aligning evaluation and professional development. Several observation systems now include "apps" or software that can be used with handheld devices and laptops to capture evidence and indicate scores during the observation. However, when using technology, the teacher and support provider or evaluator still must collaboratively develop a professional growth plan rather than relying on an automatically generated list of recommended professional development. In addition, it is necessary to have a discussion focused on which professional growth options and opportunities are most likely to help the teacher grow professionally if teaching and learning are to be positively impacted. Still, technology may expand the learning opportunities for teachers by collecting information more quickly and producing results more efficiently (see Table 2 on next page). It is important that evaluators be trained on using technological tools for evaluation during the calibration process in order to become reliable raters of teacher effectiveness.

Video technology can also be of great value because it results in records of teaching that allow for repeated viewings and discussions, and allows trained off-site evaluators to "look into" classrooms (see sidebar on page 20).

Software enables observers to record their rubric judgments quickly. Networked sites for teacher evaluation and professional development increase the opportunity to gather multiple, varied sources of evidence and learning materials in the data collection process. For example, teachers and evaluators can upload multimedia materials from classrooms to create an "e-portfolio" containing photographs, videos, Youtube® or Teachertube® clips, or audio recordings (Jans & Awouters, 2008). While the process of identifying evidence for inclusion in an electronic portfolio may expand teacher professional learning, teachers are likely to need training and support to develop the knowledge and skills to use such a tool effectively. Otherwise, some teachers may receive higher ratings because they are simply more adept at using the technology to show their practice to advantage.

In creating digital media as part of evaluation and professional development, it is necessary to ensure quality and confidentiality. For example, video capture requires the correct setup of well-functioning equipment. Digital recordings of classroom instruction and teacher activity outside the classroom are permanent records. Therefore, any system of evaluation and learning needs to be safeguarded with appropriate student, staff, and teacher permissions.

Table 2. Strengths and Cautions of Using Technology in Teacher Evaluation and Professional Development

Technological Medium	References	Strengths	Cautions
Video	(Seidel, Sturmer, Blomberg, Kobarg, & Schwindt, 2011; Sherin & Han, 2004; Sherin & van Es, 2005)	—Allows teachers to examine student thinking patterns and behaviors more closely —Classroom excerpts can be viewed repeatedly or edited into shorter clips as exemplars or special cases for future reference —Teachers can review videos with instructional coaches or teams to discuss how to improve teaching or address challenges.	Teachers need guidance to analyze their own or others' teaching individually or in group settings. Teachers need to know what evidence from the video was used to determine evaluation results and what those results mean.
Audio (e.g., "bug-in-ear")	(Rock, Gregg, Gable, & Zigmond, 2009) Pianta, La Paro, & Hamre, 2007)	—Enables real-time virtual coaching —Helps teachers understand how evaluation criteria are applied in the classroom	May cause cognitive overload for teacher in the course of teaching (attending to the coach's message and the students' needs simultaneously)
Learning Platforms	(Learning Sciences International, 2011)	—Efficiently stores evaluation results —Creates automated links to professional learning options	Instructional conversation around teaching and learning needs to be mediated by evaluator, not technology.
Digital Portfolios	(Jans & Awouters, 2008)	Allows for multiple, varied sources of evidence of teaching practice and learning materials	Teachers need to be trained in selecting standards-based evidence of teaching practice and in submitting it to the portfolio.

Technology may enhance the analysis of teacher evaluation results to support decisions made about professional development. Computers and programs make it easier to track a teacher's practice over time. In addition to individual-level results, databases can be used to analyze evaluation results at the aggregate level. For example, using the CLASS protocol (Pianta, La Paro, & Hamre, 2007), districts can create reports at the domain level to gauge teacher needs for learning opportunities related to classroom management. Therefore, instructional leaders can adapt more immediately to the professional learning needs of teachers in their district. Equally important is the capacity to analyze student learning in similar ways through analysis of test data or student growth measures. Both evaluators and teachers need to be trained in using technological tools in teacher evaluation and professional development.

VIDEO-ENHANCED TEACHER EVALUATION AND LEARNING

Recently, video has become an increasingly researched medium for teacher evaluation (Dillon, 2010). Previously, video had been used frequently in teacher education programs and somewhat frequently in teacher professional development strategies, such as "lesson study" (Abell et al., 1996; Lewis, Perry, & Murata, 2006). Video makes teachers' classrooms accessible in ways that they were not before, enabling evaluators and teachers to fully attend to the events of daily teaching (LeFevre, 2004). Recent research has supported the potential use of video individually for teachers to reflect on their own practice (Seidel et al., 2011). Seidel et al. (2011) found that teachers were engaged in viewing and analyzing video of their own practice through increased activation of prior knowledge, though they may need additional support to be more constructively critical of their own practice. It is also helpful for teachers to view video of other teachers' practices (Sherin & Han, 2004), validating the use of video for group professional development.

In particular, video analysis benefits teachers in the following ways (Seago, 2004; Sherin & van Es, 2005):

- Presents a more complex view of instruction
- Supports new norms of professional discourse
- Offers better understanding of content knowledge related to teaching (e.g., student misconceptions, communicating math more clearly)
- Improves understanding of student reasoning processes
- Refocuses conversations on student learning rather than teacher actions alone

For video to support teacher professional growth, the video needs to authentically represent teachers' practice. Evaluators and teachers must select the segments to be analyzed and discussed with a clear purpose, and provide a framework for discussion. Preliminary evidence has shown that "video clubs" influenced teachers' instructional pacing and the types of questions they ask students (van Es & Sherin, 2008). The alignment of teacher evaluation and professional development can inform research and development of video in promoting teacher effectiveness.

6. High-Quality Standards for Professional Learning

Not all professional learning opportunities are equal in quality and benefit to teachers and schools in their particular contexts. A set of professional learning standards can provide an objective appraisal of the design and implementation of teacher professional learning. Learning Forward, "an international membership association of learning educators focused on increasing student achievement through more effective professional learning" (**http://www.learningforward.org**), has defined seven Standards for Professional Learning leading to high-quality professional learning opportunities and positive outcomes for teachers and students: (1) learning communities, (2) leadership, (3) resources, (4) data, (5) learning designs, (6) implementation, and (7) outcomes. Learning Forward's standards fit well with the vision of an aligned teacher evaluation system considered throughout this brief.

High-quality standards for professional learning can help answer the following types of questions: How do you assess a school's professional learning community? What type and amount of resources are necessary to support teacher learning about student achievement data? How was a professional learning opportunity informed by a design based on research-based principles of learning? Standards for professional learning can be thought of as "book ends" for standards for teaching. Each set of standards provides a strong beginning and conclusion to an aligned system of teacher evaluation and professional development.

The successful design and implementation of an aligned system must include trained individuals to manage the system and make recommendations for professional learning opportunities that are aligned with the teachers' and district's needs and professional learning standards.

CONCLUSION

Finding ways of accurately determining teacher effectiveness is a top priority for many states and districts. The urgency of the need has pushed ahead of research on the subject, and states and districts are now attempting to find a balance between moving forward quickly but also fairly. We are offering an informal framework that may help ensure that teachers, and therefore students, benefit from the creation of aligned evaluation/professional development systems where evidence of teacher practice is used to make decisions about teacher professional growth in addition to providing information on the quality of teaching in a given classroom, school, or district. Through careful planning now, states and districts can use teacher evaluation to support dual goals—evaluation for accountability and evaluation for professional development.

Each component discussed above serves a critical role in an aligned system. Standards codify what is valued in teaching and provide a platform for common language and understanding. Clear standards provide both the guide and goal of a well-aligned evaluation system. Standards-based multiple measures of effectiveness allow for a complete picture of instructional practice, giving evidence of progress toward the goals. Multiple measures also show whether or not scores appear to be in sync. Eventually, multiple measures will also help districts make decisions about the validity of their instruments as they see consistent scoring patterns over time.

To ensure that the hard work of creating meaningful standards and selecting instruments has the intended impact, it is imperative that during the design stages district leaders make plans and identify resources to support high-quality professional development on the standards and measures at every level in the district. Everyone involved in evaluating or being evaluated should be familiar with the standards, instruments, scoring processes, types of evidence, and levels of performance. Without the key participants having this knowledge, schools and districts are left vulnerable to misunderstanding, misuse, and rejection of the new system. Good training enables shared expectations and deepens stakeholder investment.

It is worth noting that there are additional benefits to be gained from an aligned system. Both goals of teacher evaluation rely on the same foundation—sound data about classroom practice. Using these data in multiple ways is a more efficient use of resources because efforts are not being duplicated to meet both ends. Further, evidence of top performance is as valuable as evidence to the contrary. Every district has high-performing teachers who can help inform the professional development of lower performing teachers. As highlighted above, the research suggests that teachers working collaboratively and learning from each other has a stronger and longer lasting impact than having teachers take classes or bringing in outside experts for assistance.

Another benefit of an aligned system, especially one that is purposely transparent and includes teachers in every stage of development, is that it will receive greater buy-in from teachers. If teachers believe a system will be used strictly for accountability, then it will be received as punitive and less likely to be accepted. On the other hand, if teachers understand that the key role of the evaluation system is to improve teaching and learning, they can take an active role in their own development. If they understand what is expected of them, are able to spend time in professional learning activities, and can see their own practice improve, they will experience the payoff of their efforts and know that it was time well spent. As instruction improves as a whole, we will begin to see teaching elevate as a profession and teachers and schools receiving the trust, resources, and support they need to sustain excellence in U.S. education. If teachers can see these possibilities, they will carry the system forward and achieve the ultimate goal—to improve teaching and learning.

APPENDIX

Questions to Support Alignment

- **Identifying Sources of Evidence**
  - How well defined are our standards of teacher quality?
  - Do teachers understand these sources of evidence?
  - Do evaluation rubrics clearly differentiate among at least four levels of performance (e.g., Exemplary, Proficient, Needs Improvement, Unacceptable)?
  - How well aligned are evaluation instruments with the standards?
  - How comprehensive are the data these sources of evidence provide?
  - Are these sources of evidence adjustable to local contexts?
  - Do these sources of evidence apply across grade levels, subjects, and teachers of special student populations?
- **Gathering Evidence**
  - What mechanisms will inform teachers and leaders of appropriate professional development when student growth targets are not reached or when teacher observation scores are low in a particular area?
  - How will information technology be used to store, retrieve, and analyze data from teacher evaluations?
  - Who gathers the evidence of teacher effectiveness and with whom is it shared?
  - What are state and local requirements for evaluating teacher effectiveness, and what evidence is needed to meet them?
  - What are the data system requirements for capturing and analyzing the performance data of individual teachers and groups of teachers (e.g., by school, grade level, content area)?
  - Can the evaluation system help identify professional development needs identified at the classroom, grade, school, or district level?
  - How will teacher confidentiality be protected in an aligned system?
- **Responding to Evidence**
  - What resources are available at the school level to teachers to meet or exceed each of the teaching standards?
  - How will technology be used to facilitate teacher collaboration within and across schools?
  - How will changes in teacher practice and student learning be documented?
  - How does the evidence inform evaluation of resource allocation for teacher evaluation and professional development?
  - Are expectations for teacher collaboration included in teacher preparation requirements and professional standards?
  - How do teacher preparation programs support teachers in professional learning and evaluation?

REFERENCES

Abell, S. K., Cennamo, K. S., Anderson, M. A., Bryan, L. A., Campbell, L. M., & Hug, J. W. (1996). Integrated media classroom cases in elementary science teacher education. *The Journal of Computers in Mathematics and Science, 155*(1), 137–151.

Allen, J. P., Pianta, R., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and science achievement. *Science, 333*(6045), 1034–1037.

Archibald, S., Coggshall, J., Croft, A., & Goe, L. (2011). *High-quality professional development for all teachers: Effectively allocating resources*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved March 19, 2012, from *http://www.tqsource.org/publications/HighQualityProfessionalDevelopment.pdf*

Bill & Melinda Gates Foundation. (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching (MET) Project.* Seattle, WA: Author. Retrieved March 19, 2012, from *http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf*

Council of Chief State School Officers. (2011). *Interstate Teacher Assessment and Support Consortium (InTASC) model core teaching standards: A resource for state dialogue.* Washington, DC: Interstate Teacher Assessment and Support Consortium. Retrieved March 19, 2012, from *http://www.ccsso.org/Documents/2011/InTASC_Model_Core_Teaching_Standards_2011.pdf*

Croft, A., Coggshall, J., Dolan, M., Powers, E., & Killion, J. (2010). *Job-Embedded Professional Development: What it is, who is responsible, and how to get it done well* (Issue Brief). Washington, DC, and Oxford, OH: National Comprehensive Center for Teacher Quality, Mid-Atlantic Comprehensive Center, and National Staff Development Council. Retrieved March 19, 2012, from *http://www.tqsource.org/publications/JEPD%20Issue%20Brief.pdf*

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: ASCD.

Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35–39.

Denner, P. R., Salzman, S. A., & Bangert, A. W. (2001). Linking teacher assessment to student performance: A benchmarking, generalizability, and validity study of the use of teacher work samples. *Journal of Personnel Evaluation in Education, 15*(4), 287–307.

Dillon, S. (2010, Dec. 3). Teacher ratings get new look, pushed by a rich watcher. *The New York Times.* Retrieved from *http://www.nytimes.com/2010/12/04/education/04teacher.html*

Fullan, M., Hill, P. W., & Crevola, C. (2006). *Breakthrough.* Thousand Oaks, CA: Corwin Press.

Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., & Wayman, J. C. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved March 19, 2012, from *http://ies.ed.gov/ncee/wwc/pdf/practice_guides/dddm_pg_092909.pdf*

Holdheide, L., Goe, L., Croft, A., & Reschly, D. (2010). *Challenges in evaluating special education teachers and English language learner specialists.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved March 19, 2012, from *http://www.tqsource.org/publications/July2010Brief.pdf*

Jans, R., & Awouters, V. (2008). Digital portfolio: A strategy for teachers professional development. *International Journal of Emerging Technologies in Learning, 3*(1), 19–20.

Joyce, B., & Showers, B. (1988). *Student achievement through staff development* (2nd ed.). New York: Longman Publishers.

Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains* (MET Project Policy & Practice Brief). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from *http://metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf*

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2010). *Identifying effective classroom practices using student achievement data* (Working Paper, p. 51). Cambridge, MA: National Bureau of Economic Research.

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Evaluating teacher effectiveness: Can classroom observations identify practices that raise achievement? *Education Next, 11*(3).

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the prekindergarten year. *The Elementary School Journal, 104*(5), 409–426.

Learning Sciences International. (2011). *iObservation.* Retrieved from *http://www.iobservation.com/*

LeFevre, D. M. (2004). Designing for teacher learning: Video-based curriculum design. In J. Brophy (Ed.), *Using video in teacher education: Advances in research on teaching* (pp. 235–258). Amsterdam: Elsevier.

Leo, S. F., & Lachlan-Haché, L. (2012). *Creating summative educator effectiveness scores: Approaches to combining measures.* Washington, DC: American Institutes for Research.

Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher, 35*(3), 3–14.

Little, J. W. (2006). *Professional community and professional development in the learning-centered school* (NEA Best Practices Working Paper Series). Berkeley, CA: University of California, Berkeley.

Matsumura, L. C., & Pascal, J. (2003). *Teachers' assignments and student work: Opening a window on classroom practice* (CSE Report 602). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from *http://www.cse.ucla.edu/products/Reports/R602.pdf*

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33–53.

Peine, J. (2008). *The educator's professional growth plan: A process for developing staff and improving instruction* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Peterson, K. D., Wahlquist, C., & Bone, K. (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 135–153.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2007). *Classroom Assessment Scoring System.* Baltimore: Brookes Publishing.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Rock, M. L., Gregg, M., Gable, R. A., & Zigmond, N. P. (2009). Virtual coaching for novice teachers. *Phi Delta Kappan, 91*(2), 36–41.

Rock, T. C., & Wilson, C. (2005). Improving teaching through lesson study. *Teacher Education Quarterly, 32*(1), 77–92.

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research, 81*(4), 493–529.

Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation.* Chicago: Consortium on Chicago School Research at the University of Chicago.

Schalock, H. D. (1998). Student progress in learning: Teacher responsibility, accountability, and reality. *Journal of Personnel Evaluation in Education, 12*(3), 237–246.

Seago, N. (2004). Using video as an object of inquiry for mathematics teaching and learning. In J. Brophy (Ed.), *Using video in teacher education: Advances in research on teaching* (Vol. 10, pp. 259–286). Amsterdam: Elsevier.

Seidel, T., Sturmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education, 27*, 259–267.

Sherin, M., & Han, S. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education*, 20, 163–183.

Sherin, M., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*(3), 475–491.

Taylor, E. S., & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (NBER Working Paper No. 16877). Cambridge, MA: National Bureau of Economic Research.

The Danielson Group. (2011). *Framework for teaching.* Retrieved May 13, 2011, from *http://charlottedanielson.com/theframeteach.htm*

van Es, E., & Sherin, M. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education, 24*, 244–276.

Veen, K. v., Zwart, R., & Meirink, J. (2011). What makes teacher professional development effective? A literature review. In M. Kooy & K. v. Veen (Eds.), *Teacher learning that matters: International perspectives* (pp. 3–21). New York: Routledge.

Wilkerson, D. J., Manatt, R. P., Rogers, M. A., & Maughan, R. (2000). Validation of student, principal and self-ratings in 360 degree feedback® for teacher evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 179–192.

Yoon, K. S., Duncan, T., Lee, S. W., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007-No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from *http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf*

ABOUT THE NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

The National Comprehensive Center for Teacher Quality (TQ Center) was created to serve as the national resource to which the regional comprehensive centers, states, and other education stakeholders turn for strengthening the quality of teaching—especially in high-poverty, low-performing, and hard-to-staff schools—and for finding guidance in addressing specific needs, thereby ensuring that highly qualified teachers are serving students with special needs.

The TQ Center is funded by the U.S. Department of Education and is a collaborative effort of ETS; Learning Point Associates, an affiliate of American Institues For Reserach; and Vanderbilt University. Integral to the TQ Center's charge is the provision of timely and relevant resources to build the capacity of regional comprehensive centers and states to effectively implement state policy and practice by ensuring that all teachers meet the federal teacher requirements of the current provisions of the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act.

The TQ Center is part of the U.S. Department of Education's Comprehensive Centers program, which includes 16 regional comprehensive centers that provide technical assistance to states within a specified boundary and five content centers that provide expert assistance to benefit states and districts nationwide on key issues related to current provisions of ESEA.

NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
877.322.8700 | 202.223.6690

www.tqsource.org

This work was originally produced in whole or in part by the National Comprehensive Center for Teacher Quality with funds from the U.S. Department of Education under cooperative agreement number S283B050051. The content does not necessarily reflect the position or policy of the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by the federal government.

The National Comprehensive Center for Teacher Quality is a collaborative effort of ETS; Learning Point Associates, an affiliate of American Institutes for Research; and Vanderbilt University.

1821_04/12

LEARNING POINT Associates®
An Affiliate of American Institutes for Research®