



A CONTINUOUS IMPROVEMENT MODEL FOR TEACHER DEVELOPMENT AND EVALUATION

INTRODUCTION

A Rationale for Change

Today our country faces unprecedented challenges—from a “great recession” and increased global warming, to soaring healthcare costs and declining retirement security. These complex issues require an informed citizenry to debate and decide our future. But our current industrial model of education is not adequately preparing our students to successfully meet the challenges of a knowledge economy. Students must be exposed to a rich and rigorous education. They must master deeper core content knowledge than ever before as well as acquire a new set of skills to deal with the increasingly interconnected and technological world. Unfortunately, the mission of public education and consequently, the curriculum and the role of teachers, have been narrowed by those who would insist on making student performance on standardized tests the hallmark of effectiveness for schools and teachers.

Now is the moment when teachers’ voices can make a difference. We have the opportunity to seize the reins and make teaching the profession it ought to be. This is a moment of both great challenge and great opportunity. For too long, those with little knowledge about the profession and often no experience teaching, have defined for teachers what good teaching practice looks like and how it should be measured. Neither teachers nor their unions have significantly shaped the debate. In fact, in the last ten years “accountability” has been used as an excuse to exclude teachers in the discussion of how and what they should teach. But if we are to really improve practice and student outcomes, teacher involvement makes the critical difference.

Unfortunately the current discussion is centered on increasing quality by getting rid of the “bad apples,” instead of on what it takes to nurture teacher talent, foster good teaching and improve student learning. If all we do is dismiss bad teachers, we will not significantly improve public education overall. You can neither hire nor fire your way to better schools. Like other professionals, teachers continuously develop in an effort to deepen their knowledge and develop their skills. Moreover, student learning is not simply influenced by *one* teacher, no matter how great; therefore, we must think bigger and more systematically about developing teachers’ knowledge and skills, building career ladders and greater professionalism and improving the teaching and learning environments in which they work.

Teaching is complex; there is no single pedagogy that can meet the needs of every learner. Teachers bring to the classroom varying skills and knowledge that are a reflection of their training and experience. The evaluation process must reflect the complexity of teaching and learning. A system focused on truly improving practice and promoting student learning not only creates procedures for assessing individual teachers’ knowledge and skills, but also has systems of supports that provide for the continuous improvement of all teachers—high-quality supports such as job-embedded professional development, mentoring and induction programs.

Teacher evaluation must first and foremost be designed to foster teacher professional growth and student learning. While good teaching is more than just a student test score, teaching *is*, at the end of the day, about

students. We cannot ignore the importance of student learning and other student outcomes, such as developing habits and behaviors that lead to success in school and life. A good teacher evaluation system measures teachers on the practices that, over time, produce desirable student outcomes and provides teachers the opportunity to hone effective practices.

Teacher Evaluation Should Reflect the Mission of Public Schooling.

We should measure what we value, but all too often, in education we value what is most easily measured and overlook critical elements that are essential to effective practice. For example, in the current educational policy context, we focus on the percentage of students who score proficient on their state standardized tests without a critical discussion about the content standards themselves or what it means to be “proficient” in a content area. We face this dilemma when designing a teacher evaluation system. It is easy and cheap to design teacher evaluation systems that rely predominantly or exclusively on student test scores. But the consequences of doing so are detrimental to students and teachers alike. Evaluating teachers mainly on outcomes such as standardized test scores has been shown to lead to excessive test preparation, narrowing of the curriculum and focusing attention on students who are on the cusp of “proficiency” as measured by standardized tests. All students deserve a well-rounded, content-rich curriculum which encourages them to develop critical thinking skills.

What skills and competencies do we really want students to master? If our students are to compete in the 21st-century global economy and be stewards of our democracy and our planet, then we need to ensure that teachers foster students’ personal, civic and social concern and responsibility along with academic skills. These student outcomes are equally important to a vibrant democratic society, and we must find a way to measure them regardless of the challenges associated with doing so. Designing a comprehensive teacher evaluation system allows us to make sure we consider outcomes that parents, teachers and other community members care about—in addition to student test scores. We fail our students if performing well on a standardized test becomes *the* mission of public education.

The Union as a Vehicle for Change.

Teachers unions have a responsibility to teachers, the teaching profession and students to ensure that all teachers meet high professional standards of practice. For this reason, unions must play a crucial role in assessing and supporting the quality of those who enter and remain in the profession. Allowing ineffective teachers to remain in the classroom is detrimental to students, teachers, unions and the profession.

No other organization is better positioned to advance the issue of improving teaching than the AFT and its affiliates. First, we are a repository of expertise and knowledge about what good teaching practice looks like and how best to facilitate student learning. Second, we have the power to collaborate on a district level with key stakeholders to craft policies that will facilitate good teaching. And finally, teachers unions, through negotiations and legislative efforts, can institutionalize reforms at both the local and state levels to improve teaching quality and public education overall.

But teachers and their unions cannot do this alone. Comprehensive teacher evaluation must foster collective responsibility and accountability, and there must be a willingness and a readiness of all stakeholders—union leaders, administrators, policymakers, parents and the broader community—to work together. And there must be real consequences for those who are not ready or unwilling to collaborate.

Reforms to the evaluation process will only be developed and sustained with strong labor-management relationships. Collaboration, partnerships and trust at all levels are essential. Teachers and administrators

must form a true partnership to come together around the common goal of improving student learning. This will necessitate not only changing how unions and management interact, but also changing how schools are organized and governed, and how school staff members work with one another.¹

The Purpose of This Paper

This paper lays out a vision of a “smart” teacher development and evaluation framework—one that continually improves (1) as new and better research becomes available about teacher development and evaluation; (2) as data (including student outcomes) at the classroom, school and district levels are generated and analyzed; and (3) as the district builds its capacity to increase teacher quality.

Strategic planning around teacher evaluation must address how to build the capacity of school systems and school personnel to implement all the components of a comprehensive teacher evaluation system whose purpose is to grow and develop teachers to improve student learning.

The AFT’s framework is based on 10 critical principles that we believe must guide the design and implementation of any teacher development and evaluation system. Such systems must:

1. Have as their primary goals strengthening the individual and collective practices of teachers and schools to improve student learning;
2. Be developed and implemented collaboratively with teachers, not imposed on them;
3. Focus on providing continuous professional development and growth for teachers by addressing the skills, knowledge and needs of teachers depending on where they are on a career continuum (e.g., novice, midcareer, veteran);
4. Promote teacher leadership;
5. Be both formative and summative;
6. Be based on a set of standards of practice that takes into account the complexities of teaching;
7. Include evidence of teaching and student learning from multiple sources;
8. Address how to build the capacity of districts and schools to implement high-quality teacher development and evaluation systems;
9. Consider the context in which teaching and learning takes place; and
10. Be subject to continuous updating of instruments and processes as research on practice leading to valued student outcomes becomes available.

Further, it is essential that districts and their unions measure to what extent the evaluation system is being implemented with fidelity, to what extent it meets the purposes for which it was designed, and in what ways the evaluation influences teaching and learning. Collecting data on these questions will allow a district that has adopted a comprehensive teacher development and evaluation system to monitor, adjust and improve the system as necessary.

The Organization of the Framework

The framework is organized into three main sections.

- First, we distinguish between accountability for school success and individual accountability.
- Next, we discuss the purposes of teacher evaluation, and distinguish between evaluation for continuous development and evaluation for decision-making.
- Finally, we identify the components of a comprehensive teacher development and evaluation system that supports both continuous improvement and personnel decision-making.

FROM INDIVIDUAL TO SHARED ACCOUNTABILITY AND RESPONSIBILITY

Teachers are an important part of a child's education. While one teacher may be particularly significant to a student's academic success, it is the collection of teachers that a student encounters throughout his or her educational career that provides the foundation for academic success along with parents, other school staff (such as nurses, cafeteria workers, bus drivers and janitors), administrators and community members. The education community as a whole and all of its actors are responsible for providing every student with the opportunity to learn and to thrive academically, socially and emotionally.

Good teachers thrive in schools that identify, develop and support their talent. In a poor teaching environment, good teachers can and do languish. Similarly, students with good teachers may underachieve or perform poorly when their nonacademic needs are not met or when family problems interfere with learning. While assigning blame for poor student performance to an individual teacher, administrator or parent is easy and convenient, it is also counterproductive. School systems need to give parents the supports they need to nurture their children's learning and to improve system quality, not simply teacher quality. Evaluation of individual teachers has an important place in shared accountability and responsibility for student success, but it is not the only component needed to improve system quality so that all students succeed.

Implicit in the idea of shared accountability is shared responsibility and trust. Teachers must trust that administrators are true partners in the work of helping them become the best teachers they can be to successfully prepare students for life and the world of work.

PURPOSES OF EVALUATION

Prior to designing a teacher evaluation system, the most important question to answer is: Why evaluate teachers? The purposes of evaluation drive the design of the system. An effective teacher evaluation can serve multiple purposes. It can be used to:

1. Improve the overall quality of the teacher workforce by identifying and building upon individual and collective teacher strengths, and by improving instruction and other teacher practices to improve student learning;
2. Identify exemplary teachers who might serve as mentor and/or master teachers;
3. Identify ineffective teachers and develop a system of support to remediate their skills; and
4. Ensure fair and valid employment decisions, including decisions about rehiring, dismissal, career paths and tenure.

To accomplish these multiple purposes, two types of evaluation are necessary: formative and summative evaluation.

- *Formative evaluation* supports continuous growth and development. It provides teachers with feedback on how to improve their craft to promote student learning. It is a critical component of career professional growth. Data from formative evaluation also can identify specific professional

development opportunities for teachers that will facilitate student learning (e.g., instructional techniques that meet the needs of diverse learners, effective classroom management strategies, and use of student assessments).

- *Summative evaluation* is used to judge whether a standard has been met. It is used for tenure decisions, intensive assistance decisions, dismissal decisions, career path decisions and compensation decisions.

To ensure due process, formative evaluation, with opportunity for improvement, must always precede a summative evaluation.

COMPONENTS OF A TEACHER EVALUATION SYSTEM

There are five components of a comprehensive development and evaluation system that meet the formative and summative purposes of evaluation. As part of this framework, we lay out the standards for each component. Included are professional teaching standards as well as standards for measuring teaching practice (including student learning), implementation, professional contexts, and systems of support.

1. Professional Teaching Standards

Professional teaching standards advance a common, comprehensive vision of the profession. They communicate a shared understanding about what is important for teachers to know and be able to do to promote student learning and professional growth. Professional teaching standards are key to defining the practices that good teachers use to facilitate student learning. Professional teaching standards should take into account the importance of pedagogical content knowledge. Good teaching requires more than just knowing the required content knowledge. It also requires knowing how to effectively teach subject specific content to students.²

Professional teaching standards should be developed by teachers in collaboration with other stakeholders.³ Teachers not only must be familiar with the standards by which their practice is evaluated, but also must understand how those standards are applied (e.g., what excellent, acceptable and less-than-competent performance on a standard looks like).

Professional teaching standards are essential to the development of high-quality, performance-based assessments (see Box 1 for examples of professional teaching standards). In addition, professional teaching standards can be the centerpiece for professional learning communities, encouraging teachers to reflect on their practice and share that knowledge with their colleagues, thereby fundamentally reshaping the culture of schooling and reducing the isolation of teachers.

Professional teaching standards must:

- Communicate a shared vision of good teaching practice;
- Promote teacher leadership;
- Identify indicators or competencies of teacher performance;
- Address the complexity of teaching⁴ and student learning;
- Encourage teachers to be reflective practitioners;
- Attend to the collaborative and reflective nature of teachers' work; and
- Include all classroom teachers across all grades and subjects (e.g., they should reflect differences for various content areas and specialists).

Box 1. Examples of Professional Teaching Standards

- Charlotte Danielson’s Framework for Teaching
- North Carolina Professional Teaching Standards
- National Board for Professional Teaching Standards

2. Standards for Assessing Teacher Practice

An evaluation system focused on improving teaching and learning must include evidence of both good teaching practice and student learning; the system must consider the weight it gives to the evidence of each—teaching practice *and* student learning. For example, an evaluation system that values standardized student test scores more than demonstration of good teaching practice will result in undesirable teaching practices, such as narrowing the curriculum to tested subjects only or excessive test preparation. Conversely, a system that does not focus sufficiently on student learning cannot make crucial connections between teaching and learning.

Assessing Teacher Practice

When assessing teacher practice, it is necessary to:

- Use valid and reliable measures;
- Include performance standards that are appropriate for the developmental level of the teacher—novice, midcareer and veteran; and
- Incorporate a variety of evaluation techniques to capture the breadth of good teaching and professional practice, including classroom observations, review of lesson plans, self-assessments, teaching artifacts and portfolio assessments.

The Role of Student Learning and Other Student Outcomes

Student learning is at the heart of the teaching profession and must be included in any credible teacher evaluation. Several principles should guide the use of student learning in teacher evaluation:

- *Student learning should include evidence of growth in knowledge and skills based on multiple measures.* Just as no one measure can evaluate teacher performance, no one measure can or should account for student learning. Examples of evidence of student learning are: student written work, performances, group work or presentations scored using a rubric, writing samples, scores on locally designed assessments, student learning objectives (see Box 2) or student “capstone” projects (e.g., graduation, end-of-course, research or thesis paper, portfolios of art work). Teachers should document students’ progress toward mastery—their breakthroughs along with their struggles with concepts and skills—using both informal and formal tools such as written observations, surveys, rubrics, task charts, self-reflections, teacher-student conferences and individualized education or learning plans. Progress in student learning is an indicator of student success that both teachers and parents value.
- *Progress on standardized test scores may be considered as part of an overall evaluation system when the measures are valid and reliable. But progress on standardized test scores must not be the single or*

*predominant measure of student learning.*⁵ A number of practical and methodological issues exist, including:

- Many tests do a poor job of measuring student performance for both high-achieving and low-achieving students because they test a very narrow range of student learning;
- Standardized student tests that are currently in use have never been validated to measure teacher effectiveness;⁶
- Various forms of measurement error are inherent in value added and other measures of growth, and the results must be used with statistical measures of accuracy such as confidence intervals; and
- Value added is susceptible to error due to the inherent limitations of tests for capturing the complexity and breadth of student learning and the limitations of value-added methodologies.⁷

Further, where value-added models are used to measure growth in standardized test scores, they must use “best practices” in (1) the design of the model, (2) the assessment of its validity and reliability, and (3) the use of the outcomes.

Given these limitations, we recommend that progress on test scores be used only in context with a great deal of other information about teachers. We also recommend that value-added data be interpreted with expert help, used for formative purposes in teacher evaluation, and used to address whole-school accountability issues.

- *Other student outcomes, in addition to student achievement, matter.* Outcomes such as attendance, persistence and engagement have a relationship to student learning. Further, outcomes demonstrating that students have acquired habits and behaviors which lead to success in school and in life are important. Are students able to interact appropriately with their peers and adults? Can they collaborate? Can they demonstrate empathy for others? These skills are crucial to a thriving democracy and economy and should be considered in any teacher evaluation system.

Box 2. Student Learning Objectives⁸

What Are Student Learning Objectives (SLO)?

One way to measure student learning is for teachers to create rigorous student learning objectives. Student learning objectives are annual targets for growth that a teacher sets at the beginning of the year and strives to attain by the end of the year (or at the end of a semester if appropriate). They are based on a student needs assessment and aligned to the school’s standards and curricula. In addition, SLOs must be (1) based on outcomes (not activities), (2) rigorous, and (3) measurable.

Unlike standardized test scores, student learning objectives can be developed for any teacher in any subject area or grade level. Baseline data can be collected to inform student growth targets. Assessments can be standardized or teacher-developed.

The successful use of student learning objectives requires that teachers have resources and supports. Teachers need ongoing, dedicated time to assess and monitor student learning. They also need to know how to interpret data in order to adjust instruction and/or curriculum.

Putting It All Together

Determining teacher quality requires that the diverse evidence—classroom observations, parent surveys, student test scores, other evidence of student learning, etc.—be assembled into a single system to create a profile of teacher accomplishment. Measurement must consider issues such as weighting, standard setting and overall scoring.

- **Weighting:** For example, are all evaluation criteria equally important? How much consideration should be given to classroom observation data? To parent and student surveys? To student work samples and/or test data?
- **Overall scoring:** For example, what is the intent or use of an overall score or rating? Should teachers receive a rating at all? Are scores from various standards of teacher quality averaged? Is there a floor that must be reached for each standard? Can excellence on one measure compensate for weakness on another?
- **Standard setting:** For example, what does exemplary, good, acceptable and unacceptable/needs-improvement performance on a standard look like?

Ultimately, the evaluation of teacher performance requires clear definitions of what is good enough, what is exceptional, what is unacceptable, etc.

3. Implementation Standards

Implementation standards should address the important details of evaluation, such as how teachers are involved, who evaluates them, how often evaluation takes place, how the results of the evaluation will be used, and how the results are communicated to teachers. The purposes of evaluation must be considered when answering these questions.

The guidelines below should be followed for effective teacher evaluation:

- Teachers must know the standards against which they are assessed, and what constitutes excellent, acceptable and less-than-competent performance on these standards.
- Evaluators should be peers/expert teachers, and/or administrators and self.⁹
- Formative evaluations must be conducted frequently.¹⁰
- Evaluators must have formal training and demonstrate ability to assess teaching fairly and accurately.¹¹
- Evaluators must be able to interpret the findings of an evaluation in order to assist teachers in designing high-quality differentiated professional development plans.¹²
- A process for data collection and feedback must be developed.
- Standards for student achievement data quality and use must be developed.
- Systematic communication about the evaluation must take place with teachers prior to and after the evaluation process.¹³
- Ongoing professional goals must be collaboratively developed by the teacher and evaluators as part of a formative evaluation process.
- Evaluation data must inform professional development opportunities for teachers.

4. Standards for Professional Contexts

A positive professional context is essential for good teaching practice, teacher success and satisfaction, and ultimately student learning.¹⁴ Professional context describes a school's teaching and learning conditions. Teachers and students will not thrive in an environment that is not conducive to teaching

and learning. These conditions include both physical and structural elements of schools, as well as elements that influence a school's culture and climate. Measures for assessing teaching and learning conditions should consider the following factors: time, facilities and resources, teacher empowerment, leadership, professional growth, and school climate and safety.¹⁵

- **Time** refers to the opportunities teachers have to meet the needs of their students, given school schedules, noninstructional duties, paperwork and availability (or inaccessibility) of structured avenues for collaboration with colleagues (such as common planning time, lesson study and professional learning communities).
- **Facilities and resources** refer to teachers' access to the people, materials and tools they need to support their teaching. It also refers to the extent to which their schools are well-maintained and have adequate environmental conditions (such as space, lighting and ventilation).
- **Empowerment** refers to the opportunities for teachers to develop as professionals, be recognized as instructional experts, and meaningfully participate in decisions about school policies, procedures and programs that affect them.
- **Leadership** refers to how administrators and teacher leaders collaboratively shape a shared vision for success, enhance school climate, enforce norms and recognize good teaching.
- **Professional growth** refers to the quality and frequency of teachers' formal opportunities to learn what they need to know and do to be successful with the students they teach.
- **School climate and safety** refer to the quality and character of school life.¹⁶ Do teachers and students in a school feel that they belong? Do they feel safe and supported? Do they know and value one another?

Standards for positive professional contexts that support teaching and learning should be developed collaboratively by a school's staff to reflect the conditions needed to create a supportive teaching and learning environment. *A procedure for assessing a school's professional context also must be developed, and data should guide decisions about how to improve a school's teaching and learning conditions.* Collectively, all members of the school community—school board members, superintendents, principals, teachers, and other school and district staff—are responsible for ensuring that a school's teaching and learning conditions promote student academic success, and teachers' and students' social and emotional well-being.

Box 3. Sample Standards for Professional Contexts¹⁷

Standards for Working Conditions in North Carolina Schools

- Twenty-nine standards in five distinct areas—time, empowerment, professional development, leadership, and facilities and resources—serve as guidance for schools to understand positive working conditions. <http://www.ncptsc.org/>

Ohio School Climate Guidelines

- Nine guidelines are presented in areas such as engaging in school-community partnerships, conducting regular and thorough assessments toward continuous improvement, providing high-quality professional development for school leaders and staff, and engaging parents. <http://www.ode.state.oh.us/GD/Templates/Pages/ODE/ODEPrimary.aspx?page=2&TopicRelationID=433>

5. Standards for Systems of Support

Systems of support must be available throughout a teacher's career, from initial hiring through advancement, and must include a system whereby teachers identified as not meeting teaching standards are provided sufficient opportunity to improve their teaching.

Systems of support must:

- Provide a continuum of teacher support based on (1) a teacher's ability to meet teaching standards and (2) the career stage of the teacher (i.e., novice, midcareer and veteran teachers should receive support that reflects their knowledge and skills);¹⁸
- Be aligned with the professional teaching standards;
- Focus on teachers' work with students;
- Use and be informed by teacher evaluation data;
- Be intensive and ongoing; and
- Give teachers a say in improving the system based on regular and timely feedback.

Some examples of systems of support include:

- Ongoing, high-quality and relevant professional development (individual and whole-school)—evaluation results should inform professional development opportunities available to teachers;
- Induction;
- Mentoring/consulting;
- Professional learning communities;
- Lesson study; and
- Coaching.

Box 4. Examples of Systems of Support Embedded in Teacher Evaluation Systems

- *Cincinnati Public Schools Peer Assistance and Evaluation Program.* This program seeks to assist teachers in their first year in the Cincinnati Public Schools (CPS) in refining their teaching skills and orienting them to CPS, including its goals, curriculum and structure. The program also assists experienced teachers who demonstrate instructional deficiencies. Expert peer teachers work with these teachers to improve their instructional skills and bring them to the proficient level of performance as defined by CPS professional teaching standards.
- *Minneapolis Public Schools Professional Development Process.* The professional development continuum for teachers depends on systemic support, beginning with initial training and collaboratively supported practices, evolving into independently and collegially facilitated growth, and continuing throughout the teaching career with ongoing reflection and leadership. The Minneapolis Standards of Effective Instruction apply to all teachers and assist them as they move through the professional development continuum. These standards are expected to be used as a guide toward planning and implementing staff/professional development to support teaching quality and student achievement. They are also an effective tool in coaching, mentoring and teaming as a part of the professional development process.

CONCLUSION

Teachers need to take responsibility for their profession, define what it means to be a good teacher, and play a role in deciding who should enter and remain in the profession. Teachers can do this by leading the effort to overhaul teacher evaluation rather than reacting to the plans of others.

For too long, we have tolerated teacher evaluation systems that are mere formalities designed to meet a state or district mandate. Worse, when evaluation is used only for punitive reasons, it does not improve practice or increase student learning. Teacher evaluation, for most teachers, has not been about access to meaningful professional development or about opportunities to discuss how to improve their schools. It has not been about ensuring quality. It has not met the unique needs of either novice or veteran teachers in critical areas such as instruction and assessment, classroom management, parental involvement and teacher collaboration. Teacher evaluation in most school districts is not the catalyst for professional growth. It is time for that to change.

Teachers know what good teaching is—it is inspiring children to explore their world, learn how it operates and how to express their understanding of it, and respect others with different opinions. It is challenging students to acquire more knowledge and use it wisely. It is providing them with the opportunity to reach their potential intellectually, socially and emotionally. Good teachers have high expectations for their students, and use a variety of materials and resources to plan lessons, monitor instruction and assess student learning. Good teachers know the value of collaborating with other teachers, parents and administrators to ensure their students are successful. Good teachers understand that teaching is not merely pouring content into children; it is about facilitating learning: motivating children to learn, giving them the support necessary to develop skills and knowledge, and helping them overcome problems and assume responsibility for their actions and their learning.

Good teachers are not born; rather, they are carefully and systematically cultivated through rigorous recruitment, preparation, induction and continuous professional development. Yes, comprehensive teacher evaluation, when done right, can weed out those who should not remain in the profession. But more important, it can take good teachers and make them *great*. Teaching is a profession built on the hard work, reflection, care, persistence and intellect of great teachers. We must do everything we can to ensure we protect the profession and provide our students with an education that will truly prepare them for the future.

¹As Stein and Matsumura (2009) contend, “these new demands for teaching will require attendant changes in schools and districts. We cannot attain ambitious goals for student learning without concurrent changes in structures and systems that support improved instruction and teacher learning” (p. 180). For further discussion, see Stein, M.K., & Matsumura, L.C. (2009). *Measuring instruction for teacher learning*. In D. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 179-205). Washington, DC: Sage Publications.

²Ball, D.L., & Hill, H. C. (2009). *Measuring teacher quality in practice*. In D. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 80-98). Washington, DC: Sage Publications.

³Research highlights the importance of including teachers in both the design and implementation of the evaluation system. For a discussion of this, see Joint Committee on Standards for Educational Evaluation (1995). *The program evaluation standards (2nd ed.)*. Thousand Oaks, CA: Sage.; and Kyriakides, L., Demetriou, D., & Charalambous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research*, 48(1), 1-20. doi:10.1080/00131880500498297

⁴Green, T.F. (1971). *The activities of teaching*. New York, NY: McGraw-Hill.; and Pacheco, A. (2009). Mapping the terrain of teacher quality. In D. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 160-178). Washington, DC: Sage Publications.

⁵Stein and Matsumura argue that “given its prominence in teacher quality discussions today, it is important to articulate why student scores on achievement tests alone cannot shoulder the burden of the improvement agenda” (p. 199). They offer several reasons for not basing teacher quality judgments solely or predominantly on student scores on achievement tests: “achievement scores are indicators ... [and] provide little insight into the mechanics of how to improve instruction.... standardized achievement tests are not designed to test the range of content and skills that students need to learn over the course of an academic year.and standardized achievement tests also tend to focus on lower-level skills” (p. 199). Further, Stein, and Matsumura contend that measuring teacher quality using student standardized test scores does not address all the skills and knowledge needed for teachers to teach and students to learn.

⁶ According to the AERA, NCME, APA (1999) standards for educational and psychological testing, all assessments should be validated for their intended purposes. For more information see American Educational Research Association, American Psychological Association, and National Council for Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Research Association.

⁷Research shows us that even the best value-added models provide measures of student learning that vary enormously from year to year, especially for individual teachers (versus whole school), and even more so for teachers in small classes and small schools. Reasons for the annual instability include:

- substantial statistical “noise” in both the pre- and post-test years from small sample size, test measurement error, sampling error, and changes in the classroom and school environment outside the teacher’s control;
- tests with inadequate floors and ceilings, thus hiding actual student progress;
- testing periods that include two teachers—but with the results attributed to only one teacher—when tests are administered before the end of the school year;
- non-random assignment of students to teachers;
- bad data systems and mismatching of students to teachers; and
- failure in some value-added models to include student background and the fade-out of prior teacher effects.

⁸Austin Independent School District. (2009). AISD REACH: Program overview. Retrieved from <http://www.austinisd.org/inside/initiatives/compensation/overview.phtml>.

⁹Research shows that teachers are more receptive to and incorporate instruction and advice from those who have significant knowledge of curriculum, instruction and academic content, such as expert teachers. For a discussion of this, see Stiggans, R. J., & Duke, D. L. (1988). *The case for commitment to teacher growth: Research on teacher evaluation*. Albany, NY: State University of New York Press.; and Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Case studies for teacher evaluation: A study of effective practices*. Santa Monica, CA: RAND Corporation. Retrieved from <http://www.rand.org/pubs/notes/2007/N2133.pdf>.

¹⁰Multiple observations are essential to a fair evaluation process as well as the improvement of practice. Current research suggests including at least four or five observations in an overall single evaluation. For a discussion of this, see Blunk, M. (2007, April). *The QMI: Results from validation and scale-building*. Paper presented at the annual meeting of the American Research Association, Chicago, IL.

¹¹Research points to the importance of systematically training classroom observers and evaluators. For a discussion of this, see Joint Committee on Standards for Educational Evaluation. *The program evaluation standards (2nd ed.)*.

Thousand Oaks, CA: Sage; and Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://www.tqsource.org/publications/practicalGuide.pdf>.

¹²Stein and Matsumura (2009) maintain that “just as teachers benefit from knowing the variety of ways in which students might arrive at a wrong answer, it would be important for assessors of teachers to be able to spot typical missteps that teachers make” (p. 189).

¹³Communication prior to, during and after the evaluation is essential. For a discussion of this, see Joint Committee on Standards; Mathers, C., Oliva, M., & Laine, S. (2008). *Improving instruction through effective teacher evaluation: Options for states and districts*. Washington, DC: National Comprehensive Center for Teacher Quality; and Wise, Darling-Hammond, McLaughlin, & Bernstein (1984).

¹⁴Various studies have linked teacher working conditions and school climate to increased student academic achievement, including: Bryk, A. S. & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage Foundation; Hirsch, E. & Emerick, S. (2007). *Teacher working conditions and student learning conditions: A report on the 2006 North Carolina Teacher Working Conditions Survey*. Hillsborough, NC: Center for Teaching Quality.; Sherbloom, S. A., Marshall, J.C., & Sherbloom, J. C. (2006). The relationship between school climate and math and reading achievement. *Journal of Research in Character Education*, 4(1&2), 19-31; and Whitlock, J. L. (2006). Youth perception of life in school: Contextual correlates of school connectedness in adolescence. *Applied Developmental Science*, 10, 13-29.

¹⁵Adapted from the Center for Teaching Quality and the New Teacher Center’s survey work on teaching and learning conditions. See Berry, B., Fuller, E., & Williams, A. (2008). *Final report on the Mississippi Project CLEAR Voice teacher working conditions survey*. Washington, DC: Center for Teaching Quality; and Hirsch, E., Freitas, C., Church, K., & Villar, A. (2008). *Massachusetts Teaching, Learning and Leading Survey: Creating Conditions Where Teachers Stay and Students Thrive*. Santa Cruz, CA: New Teacher Center.

¹⁶Cohen, J., McCabe, E. M., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice and teacher education. *Teachers College Record*, 111(1), 180-213.

¹⁷Hirsch, E. (2008). *Identifying professional contexts to support highly effective teachers*. Washington, DC: National Comprehensive Center for Teacher Quality.

¹⁸Stein and Matsumura assert that assessment ought to be educative; in other words, measures of instruction should “provide information about the next step that teachers must take to continue to improve and develop” (p. 189). These measures should “be based on articulation of a developmental trajectory of teacher learning” (p. 189). The authors argue that even though the empirical research base for this needs strengthening, we can “begin to think about ways in which measures could point the way toward what a teacher would have to do to get to the next level ... [and could also be] sensitive to how things can go wrong in practice” (p. 189).