

Getting It Right

A Comprehensive Guide to
Developing and Sustaining
Teacher Evaluation and
Support Systems

Getting It Right

A Comprehensive Guide to Developing and Sustaining
Teacher Evaluation and Support Systems



1525 Wilson Boulevard, Suite 500 · Arlington, VA 22209
www.growgreatschools.org · 1-888-780-7805

Acknowledgments

The National Board for Professional Teaching Standards (NBPTS) is pleased to release this document. The inspiration came from the Board of Directors, which wanted to provide states and districts guidance in how to effectively evaluate and support teachers. The development and writing of the document was guided by the NBPTS Certification Council and Executive Committee whose insight and expertise is reflected on every page. This document would not have been possible without the high level of commitment and support that was provided at all phases of its development.

Edited by

Joan E. Auchter, Chief Innovation Officer

Emma K. Parkerson, Project Management Specialist

Editorial Consultants

Robert Johnston, *The Hatcher Group*

Matt Coleman, *Drexler*

Copyright © 2011 by NBPTS

Contents

6	Preface
8	Overview
10	How Students Benefit
11	The Need
14	Stage 1: Developing an Evaluation and Support System
26	Stage 2: Implementing and Managing the Evaluation and Support System
29	Stage 3: Conducting Collaborative Evaluations that Enhance Teaching Practice and Student Learning
33	Stage 4: Supporting Teachers
38	Outcomes
39	Resources
41	Conclusion
42	Bibliography

Preface

To help ensure that all students are taught in a successful manner, states and school districts face a growing urgency to differentiate the effectiveness of their teachers' practices. Grounded in its mission and core competence of assessing effective teaching, the National Board for Professional Teaching Standards (NBPTS) prepared *Getting It Right* to assist states and school districts in the development and operation of their teacher evaluation and support systems.

From its earliest days through the present, NBPTS has enlisted the most creative and brightest minds in teaching, educational assessment, and psychometrics to develop cutting-edge processes, to incorporate new techniques, and to push the field of teacher assessment to new frontiers. The National Board demanded that the leading psychometricians study and review each phase for reliability, validity, bias, and adverse effects. One could argue that by involving dozens of universities, research labs, scholars, teacher leaders, and premier testing organizations around the nation, the National Board's investment in developing a voluntary, advanced certification system has built the field of teacher performance assessment to which states and districts turn today. For more information about the history of NBPTS, please see *What Teachers Should Know and Be Able to Do* or *Advances in Program Evaluation*.

In all its work, the National Board is guided by a simple premise: the hallmark of accomplished teaching is student learning. *Getting It Right* follows the basic tenets of the National Board's certification process: Teacher assessment involves multiple measures, a rich array of generated evidence, a collaborative process, and feedback that supports ongoing teacher growth.

Effective teaching with all the desired outcomes is multifaceted and challenging to measure. Using this guide will enable states and school districts to establish successful teacher evaluation and support systems. These systems will not only promote high levels of student performance, improve teaching practice, and create more effective school environments, but also strengthen the public's understanding of what teachers should know and be able to do.

Overview

“It is critical that teachers embrace and trust a new teacher evaluation system, secure in the knowledge that classroom teachers have been full partners in its design and that it truly has the potential to transform practice and dramatically increase student achievement.”

— Linda Darling-Hammond,
Charles E. Ducommun Professor of Education at Stanford University

Research tells us that teacher effectiveness is the strongest predictor of student achievement. The National Board for Professional Teaching Standards, created more than 20 years ago, is based on this premise and is charged with (a) defining what accomplished teaching looks like and (b) developing a method to effectively measure it.

Today, states and school districts are working to ensure that more classrooms are led by effective teachers. To get there, they are exploring the best ways to assess teachers and to gauge a teacher’s effect on student performance and teacher practice.

One of the most important outcomes of those efforts will be restoring credibility to teacher evaluation while providing support. Today, most evaluators give most teachers positive ratings on summative evaluations, according to the Center for American Progress, which researched the issue for its report “So Long, Lake Wobegon?”. In 2009, 60 percent of Denver Public Schools teachers were told on their evaluation that they did not need to improve in any area. Eighty-eight percent of Elgin, Illinois tenured teachers received an “excellent” rating between 2003 and 2007 (Weisberg et al, 2009). Further, as reported in *Improving Teacher Quality Around the World: The International Summit on the Teaching*

Profession, the Organization for Economic Co-operation and Development (OECD) Teaching and Learning International Survey program (TALIS), “[an] international survey of teachers shows that the vast majority of teachers welcome appraisal of and feedback on their work, and report that it improves their job satisfaction and effectiveness. However, overall, too many teachers report that they do not receive any feedback on their work.” There is general agreement among both friends and critics of public education that our teacher evaluation systems need to be drastically overhauled.

Drawing from its own experiences in certifying accomplished teachers and some of the most promising research and lessons from the field, the National Board developed this comprehensive guide as a way to make sure teachers have a voice in the support and guidance they receive to shape their growth at all stages of their career.

This guide provides a much needed resource for states and districts to use as they develop their own teacher evaluation and support systems. More than anything else, new evaluations must be conceived, developed, and implemented as part of evaluation *systems* that enhance teacher practice, influence student learning, and improve schools. The stages outlined in this guide can help shape systems that are consistent with local policies—from recognizing stellar teachers to identifying the best strategies for supporting most teachers and helping others out of the profession.

A well-functioning teacher evaluation system goes beyond the checklists commonly used in schools. The system must specify what will be measured, define how it will be measured, clarify how the measures will be applied consistently, lay out a plan for providing feedback and continuous support, and have buy-in and leadership from key stakeholders. It will also highlight how to use the evaluation results to improve the school culture, teacher practice, and student outcomes.

Designing, implementing, and managing effective evaluation systems involve integrating many distinct but related dynamics. This framework requires clear leadership, continuous collaboration with teachers and other stakeholders, and deliberate steps. Further, an effective system also sets out standards of teacher performance, defines consequences for exceeding or failing to achieve those standards, provides continuous professional development, requires detailed conversations between evaluators and teachers, and promotes a culture of learning.

How Students Benefit

Any strong evaluation system should result in students meeting their highest potential. Students benefit from

- ▶ Stronger and more consistent teacher practice
- ▶ Supportive learning environments that nurture learning
- ▶ Teachers who collaborate to meet the needs of the whole child
- ▶ Teachers who can engage students' families in the learning process
- ▶ Adults who model behaviors and attitudes for success
- ▶ Effective teachers who deliver instruction with academic rigor
- ▶ Teachers who can prepare students to flourish in a global society

Although a single evaluation tool will not produce all of those benefits, the guiding principles here will result in systems that foster maximized teacher effectiveness and student performance.

The Need

To move from relying on an isolated event as the sole indicator of teacher practice, the National Board's *Getting It Right: A Comprehensive Guide to Developing and Sustaining Teacher Evaluation and Support Systems* includes the components that the field recognizes as critical to driving effective teaching practice, retaining effective teachers, and sustaining high levels of student performance.

As the country is moving from a single student achievement score to evaluating multiple measures of student learning and growth, teacher evaluation is experiencing a similar nexus. It is important that everyone involved in developing a teacher evaluation system understand that the goal must be more than that of developing an evaluation instrument that is used at a single point in time and is not linked to student learning. Too often the measurement is seen as an end rather than a means to an end. In the case of a teacher evaluation system, the end is teacher, school, and student growth. While the need to reform teacher evaluation is real, the need to get it right is critical.

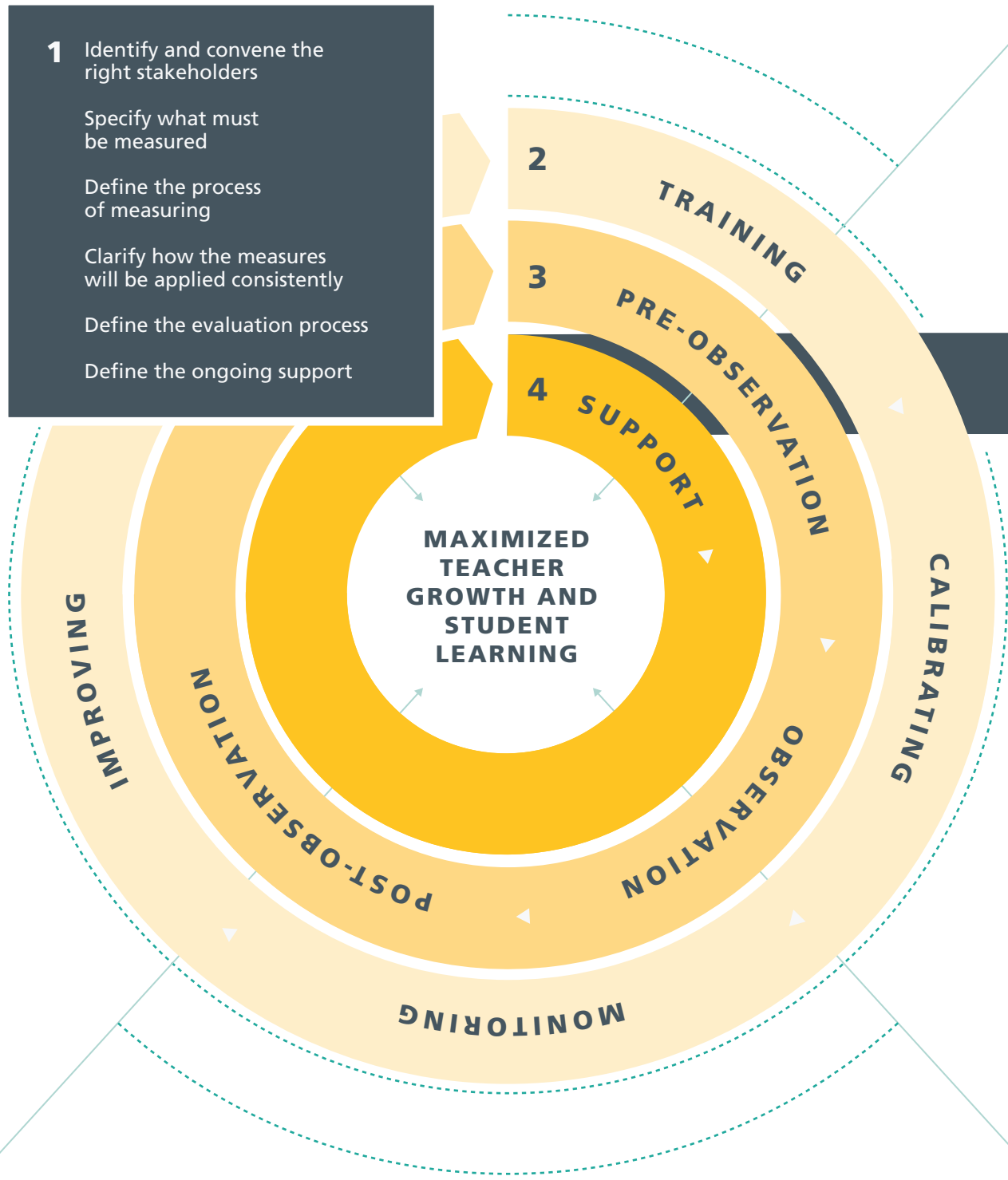
The National Board's framework for developing and sustaining a teacher evaluation and support system includes four major stages:

1. Developing an Evaluation and Support System
2. Implementing and Managing the Evaluation and Support System
3. Conducting Collaborative Evaluations that Enhance Teaching Practice and Student Learning
4. Supporting Teachers

Getting It Right

A COMPREHENSIVE GUIDE TO DEVELOPING AND SUSTAINING
TEACHER EVALUATION AND SUPPORT SYSTEMS

- 1 Identify and convene the right stakeholders
 - Specify what must be measured
 - Define the process of measuring
 - Clarify how the measures will be applied consistently
 - Define the evaluation process
 - Define the ongoing support



Stages

1

DEVELOPING THE EVALUATION & SUPPORT SYSTEM

2

IMPLEMENTING & MANAGING THE EVALUATION & SUPPORT SYSTEM

3

CONDUCTING COLLABORATIVE EVALUATIONS

4

SUPPORTING TEACHERS

OUTCOMES

- Higher student performance
 - Enhanced teacher practice
 - Sustainable school growth
 - Increased teacher retention
 - Increased student engagement
 - Collaborative learning communities
 - Improved return on investment of professional development dollars
 - Strengthened human capital in high-need schools
-

STAGE **1**

Developing an Evaluation and Support System



Developing an effective evaluation and support system involves multiple steps and multiple constituencies to ensure a sustainable foundation. This development effort must be thoughtful, must be purposeful, and must have the time and include the stakeholders requisite to building a system that results in effective teacher practice and student learning. As schools, school districts, and school boards collaboratively prepare these systems, they must also review existing regulations, guidelines, or any other contractual obligations that would inform the development process and ensure compliance.

Developing an effective teacher evaluation and support system involves six major tasks:

- ▶ Identify and convene the right stakeholders.
- ▶ Specify what must be measured.
- ▶ Define the process of measuring.
- ▶ Specify implementation plan and training systems.
- ▶ Clarify how the measures will be applied consistently.
- ▶ Define the evaluation process.
- ▶ Define the ongoing support.

TASK 1

Identify and Convene the Right Stakeholders

From the beginning, developing a teacher evaluation system must involve the right groups of people: teachers (including National Board Certified Teachers) and others who have a vested interest in enhancing teacher practice and improving student learning, including school board members, parents, and the greater community. This collaboration is critical. Although the development team will reflect the dynamics of the state or district, it should include a cross-representation of learning community members. The responsibility of the team, including teachers to ensure ongoing buy-in and credibility, must not end as the system is approved. Some subset of the team should be included in the continuous monitoring and improvement of the system.

TASK 2

Specify What Must Be Measured

This comprehensive guide uses the National Board's Five Core Propositions to specify the knowledge, skills, abilities, and commitments that all teachers must demonstrate (see *Resources*). Those propositions were established in 1989 as the cornerstone of the system of National Board Certification and continue to serve as the foundation for all National Board Standards. Moreover, they have been integrated by school districts, states, colleges, universities, and other entities to strengthen teacher preparation and teaching.

The five core propositions apply to all teachers, from classroom teachers to guidance counselors and library media specialists. Consequently, as states and school districts develop their evaluation indicators, they may use the National Board Standards and scoring rubrics to adapt the core propositions to specific job requirements and responsibilities. Grounded in the five core propositions, the National Board Standards and scoring rubrics are available for 16 specific content areas and 6 student developmental levels, which cover 95 percent of all teaching practices.

The Five Core Propositions

1. Teachers are committed to students and their learning.

Teachers recognize individual differences in their students and adjust their practice accordingly.

Teachers have an understanding of how students develop and learn.

Teachers treat students equitably.

Teachers' missions extend beyond developing the cognitive capacity of their students.

2. Teachers know the subjects they teach and how to teach those subjects to students.

Teachers appreciate how knowledge in their subjects is created, organized, and linked to other disciplines.

Teachers command specialized knowledge of how to convey a subject to students.

Teachers generate multiple pathways to knowledge.

3. Teachers are responsible for managing and monitoring student learning.

Teachers call on multiple methods to meet their goals.

Teachers orchestrate learning in group settings.

Teachers place a premium on student engagement.

Teachers regularly assess student progress.

Teachers are mindful of their principal objectives.

4. Teachers think systematically about their practice and learn from experience.

Teachers are continually making difficult choices that test their judgment.

Teachers seek the advice of others and draw on education research and scholarship to improve their practice.

5. Teachers are members of learning communities.

Teachers contribute to school effectiveness by collaborating with other professionals.

Teachers work collaboratively with parents.

Teachers take advantage of community resources.

TASK 3

Define the Process of Measuring

To be meaningful, any system designed to measure teacher performance must be calibrated so it is of maximum value to teachers and students. For teacher observations within the evaluation system, the scoring rubric is a tool that describes specific characteristics of performance at varying levels of achievement in order to clarify expectations or feedback and to limit misunderstandings in expectations (Mertler, 2011; Moskal, 2000). Consequently, these rubrics must include (a) differentiated performance levels, (b) criteria or indicators, and (c) types of evidence at different performance levels. In addition, before implementing the rubric as part of the evaluation system, the following questions must be answered: Is each criterion important and discrete? Are the levels of performance meaningful and clearly differentiable? Are the descriptors clearly defined and unambiguous, without being overly specific? Although the creation of sound and valid rubrics can be both complex and time-consuming, it is necessary for a valid and effective evaluation system.

Develop differentiated performance levels. A hallmark of the National Board Certification process is its integrated system: the core propositions are the foundation

for the standards; the assessment requires demonstration of the knowledge and skills defined in the standards; and the scoring rubric is standards- and evidence-based, thereby establishing the level of performance required.

Stiggins noted that “constructive assessment [involves] defining the achievement targets” (2001, p. 3). Differentiated performance levels are the ratings that distinguish the varying levels of quality of teacher practice. They may be expressed with qualitative titles or a quantitative point system. Some states may name their levels according to the level of independent teacher practice (e.g., emergent, supported, functional, and independent). Regardless of the terms that are selected, they should be parallel in construction and usage, making it easier for the target audiences to understand how the terms relate across a spectrum.

Two commonly used terms that are problematic are “not meeting” and “developing.” The issue with “not meeting” is its dependence on a “what” to describe what is not being met; it fits more readily with a checklist evaluation. The issue with the term “developing” is that all teachers, independent of their effectiveness, should be developing. Selecting the titles and defining the levels is a critical first step that should be done purposefully and with great care.

Teacher evaluation systems must have differentiated levels of teacher performance that are aligned with the standards of professional practice and that are clearly defined and communicated to all stakeholders. Next is one such example of qualitative levels aligned with the standards of practice.

Level	Definition as Applied to Standards of Professional Practice
Exemplary	Consistently applies knowledge, skills, and dispositions to create , manage, monitor, and adapt a classroom culture in which each student achieves and sustains the desired results
Proficient	Applies knowledge, skills, and dispositions to manage, monitor, and adapt classroom strategies so that each student achieves and sustains the desired results
Adequate	Applies selected knowledge, skills, and dispositions to implement classroom strategies so that most students achieve the desired results
Ineffective	Applies limited knowledge, skills, and dispositions to inconsistently implement classroom strategies so that some students achieve undesired results

Develop sample rubrics. If one is to provide teachers with a clear and common understanding of the evaluation targets, differentiated performance levels must be accompanied by exemplars or clear indicators, and they should contain consistent descriptions of performance criteria as well as explicitly stated attributes. Within a formative context, teachers who use such rubrics then have an opportunity to build on their initial performance and to adjust their learning accordingly. Rubrics become tools in the evaluation when they are designed to promote growth in teaching practice and consistency in administration.

Each standard of professional practice must have a set of indicators that are differentiated according to the performance levels. States and school districts must develop rubrics that integrate and articulate the specific performance for each standard and each indicator at each level on the basis of the appropriate instructional framework. For those states/districts that design their own rubric, ensure time to review for validity and reliability. A first step in the process is to match the indicators to the appropriate core proposition. The example shown next focuses on the first core proposition of teacher practice and the first indicator for that proposition— “Teachers recognize individual differences in their students and adjust their practice accordingly”—and then differentiates the observable behavior for each of the four performance levels.

After developing a rubric, it is critical to have the rubric reviewed for validity and reliability, answering the questions: Is the rubric measuring what it was designed to measure and are the levels meaningful? Can evaluators apply the rubric in a standard manner?

Core Proposition 1. Teachers are committed to students and their learning. *Teachers recognize individual differences in their students and adjust their practice accordingly.*

Teachers who demonstrate exemplary practice	gather and analyze social, emotional, and cognitive data from multiple sources, including self-developed systems , about each student to inform and adapt their teaching.
Teachers who demonstrate proficient practice	gather and analyze social, emotional, and cognitive data from established sources about each student to inform and adapt their teaching.
Teachers who demonstrate adequate practice	use provided social, emotional, and cognitive data from established sources about their students to inform their teaching.
Teachers who demonstrate ineffective practice	inconsistently use provided social, emotional, and cognitive data from available sources about their students to inform their teaching.

Identify types of evidence that demonstrate performance levels. Although the performance levels and scoring rubrics used to assess teacher practice are critical to the process, they alone are not sufficient. The National Board process requires multiple evidence-based measures of teacher performance, including (a) demonstrated content knowledge, (b) videos of small-group and whole-classroom interaction, (c) student work samples, (d) teacher written documentation and reflection, and (e) other materials such as parent and community involvement that demonstrate the effect on student learning. It is critical that any teacher evaluation system identify multiple measures of teacher performance to get a full and valid picture of the teacher's effect on student learning and to avoid over-reliance on any one measure.

Examples of multiple evidence-based measures for teachers include, but are not limited to,

- ▶ Demonstrating academic rigor in all student assignments
- ▶ Lesson planning
- ▶ Using ongoing evaluations of students' work
- ▶ Mentoring other teachers
- ▶ Using current research on teaching practice in specific grade levels or subjects
- ▶ Actively participating in grade-level or subject teams to support all students
- ▶ Engaging families in their children's learning
- ▶ Using multi-year trends to show growth in student learning through measures

Student learning is always identified as one of the multiple measures. Additionally, evaluation systems should be structured to rely on evidence of student learning that is explicitly aligned with the elements of curriculum for which the teachers are responsible, thus ensuring that teachers are evaluated for whom and for what they are teaching. Multiple measures of student learning could include such evidence as

- ▶ Increased ability to ask and respond to deep questions
- ▶ Motivation and self-efficacy
- ▶ Decreased tardiness and student discipline incidents
- ▶ Increased attendance

- ▶ Examples of student work
- ▶ Growth in performance on classroom, school, and district curriculum-based evaluations
- ▶ Achievement scores on validated nationally normed assessments

When aligned and valid, student achievement scores could be one measure of student growth. The National Board's 2011 report by the Student Learning, Student Achievement Task Force differentiated student learning from student achievement. "Student achievement is the status of subject-matter knowledge, understandings, and skills at one point in time.... Student learning is growth in subject-matter knowledge, understandings, and skills over time.... It is student learning—not student achievement—that is most relevant to defining and assessing accomplished teaching" (p. 28).

The task force set out the following essential criteria that should be adopted when incorporating the use of standardized student assessments in teacher evaluation.

The assessments should

- ▶ Be aligned with the curriculum and student learning goals a specific teacher is expected to teach
 - ▶ Be constructed to evaluate student learning—that is, performance at two or more points in time, so that changes in students' understanding and performance can be substantially attributed to a teacher's instruction
 - ▶ Be sensitive to the diversity of students, including those with special needs or limited English proficiency, as well as gifted or high-achieving students
 - ▶ Capture learning validly and reliably at the students' actual achievement levels
 - ▶ Provide evidence about student performance and teacher practice that reflects the full breadth of subject-matter knowledge and skills that are valued
-

To provide an accurate representation of the teacher's role, the evaluation system or the multiple measures must account for such variables as students who begin the year significantly behind, who have transferred in during the school year, who are ready for greater challenges, who speak limited English, or students with disabilities or language-acquisition needs. Further, measures must be made at multiple points in time to track students' improvement or lack of improvement.

Additionally, the task force recommends that assessments or evaluations of teacher practice:

- ▶ Be grounded in student learning, not student achievement.
- ▶ Employ measures of student learning explicitly aligned with the elements of curriculum for which the teachers are responsible.
- ▶ Attribute student growth to the teachers responsible.
- ▶ Establish the link between student learning and teacher practice.
- ▶ Use measures that, to the greatest extent possible, reflect the full curriculum, the full scope of a teacher's responsibilities, and the full domain of skills and competencies students are expected to develop.

“While those who evaluate teachers could take student test scores over time into account, they should be fully aware of their limitations, and such scores should be only one element among many considered in teacher profiles” (Baker, et al 2010). Typically, combining value-added performance measures with individual classroom observations is the best general alternative measure of teacher practice (Harris, 2011). This practice is limited to those grades and content areas that have value-added scores. “Student test results are typically available only for grades 3 through 8 and in reading/language arts and mathematics (and, increasingly, science). Furthermore, these tests are often not scaled in a way that permits the measurement of growth from year to year. Estimates suggest that, even in states with vertically scaled tests, only about 30 percent of K-12 teachers would have such student test scores available to develop measures of student growth in achievement for teacher evaluations.” (NBPTS, 2011)

TASK 4

Clarify How the Measures Will Be Applied Consistently

It is critical to develop procedures to ensure the evaluation system is valid, reliable, fair, and meaningful. Those procedures also must entail a comprehensive communication plan to ensure that all stakeholders comprehend the components of the evaluation system and how they will be applied. A key step is the development of benchmarks that establish the rating levels and rubric and that provide the standard by which the rating levels can be consistently applied. This development is done by a permanent monitoring team representing various stakeholder groups that also develops training materials and training.

During the development stage, a team of qualified staff members—the permanent monitoring team—is charged with ensuring that the evaluations are conducted in accordance with the observation rubric standards, indicators, and rating levels, as well as with the selected observation materials. With input from teachers and other key constituents, this team develops the materials that will be used to train all evaluators to rate performances reliably and to rate bias prevention training to ensure fairness. Although interpretation of the rating levels is a policy decision, anchoring and maintaining the stability of the observations should be done over time by the same individuals who have the background knowledge and training to identify and rate effective teaching. This team should develop such materials as

- ▶ Explanations of the standards, levels, indicators, and artifacts
- ▶ Anchor observations for each level and indicator of the scoring rubric
- ▶ A certification test to measure each observer's ability to apply the scale, resulting in score-scale stability and inter-rater reliability that is based on observed practice
- ▶ Monitoring and recalibration observation materials to ensure that observers are true to the desired measures
- ▶ Measures with associated consequences
- ▶ A system for managing the data
- ▶ A process for developing professional growth plans
- ▶ Specifications for conducting difficult conversations about practice

The application of training and monitoring will be presented in the upcoming section titled *Stage 2: Implementing and Managing the Evaluation and Support System*.

TASK 5

Define the Evaluation Process

Effective teaching involves teaching practice grounded in an understanding of how to facilitate student learning that leads to student growth. The coupling of teaching and learning requires observations over time with artifacts that support the student learning. The end goal is that observation results in professional growth for every teacher.

As the process is best defined through application, the following elements of the evaluation process will be presented in the upcoming section titled *Stage 3: Conducting Collaborative Evaluations That Enhance Teaching Practice and Student Learning*:

- ▶ Define the observation process.
- ▶ Communicate results of observation levels.
- ▶ Specify the evaluation results.
- ▶ Specify the consequences.
- ▶ Specify the appeals process.

TASK 6

Define the Ongoing Support

“In a recent survey of over 6,100 Massachusetts Teacher Association (MTA) members, 76 percent indicated that evaluations identify their strengths, but only 34 percent indicated that evaluations identify professional development that would improve their practice”

(Skinner, 2010). Good observations provide good feedback, which leads to better support of instruction—resulting in student growth. At a minimum, all good support should be

- ▶ Collaborative
- ▶ Evidence based
- ▶ Timely
- ▶ Developmental
- ▶ Growth oriented

Because the process is best defined through application, the following elements of ongoing support will be presented in the upcoming section titled *Stage 4: Supporting Teachers*:

- ▶ Specify differentiated support.
- ▶ Monitor effectiveness of support.
- ▶ Highlight use of evaluation results to improve school culture, teacher practice, and student outcomes.

STAGE **2**

Implementing and Managing the Evaluation and Support System



Implementing an effective teacher evaluation and support system involves leadership and partnership at multiple levels so that the right constituencies understand the goal of the system and buy into and provide the financial and human resources to support the process. Effective training and monitoring ensure that all participants have the same understanding of the purpose and outcomes, resulting in all teachers receiving fair, equitable treatment and constructive feedback.

TASK 1

Implement the Evaluation and Support System

The initial step involves establishing the process and schedules for training and monitoring, including

- ▶ Roles and responsibilities of trainers and observers
- ▶ Schedule of training
- ▶ Frequency of observations
- ▶ Types of monitoring

Before implementation, all constituents, including teachers, should receive

- ▶ Communication related to rollout of the evaluation system
- ▶ Orientation defining the purpose, components, and all outcomes of evaluation and support and describing the system's collaborative development process
- ▶ Training in standards, indicators, levels, and processes
- ▶ An overview of the ongoing process and everyone's role, including their own
- ▶ A plan for an ongoing communication loop throughout implementation

Using the training process and materials created in the development stage, the monitoring team will promote effective application of the evaluation tool and reliability of the ratings through

- ▶ Training observers to award ratings using benchmark observations linked to the standards, indicators, and levels
- ▶ Training a subset of teachers to support the process
- ▶ Certifying observers by administering score-scale and rater reliability tests.

TASK 2

Manage the Evaluation System

The evaluation system requires continuous monitoring of the training, observation ratings, and systems to inform the type of support and development that would ultimately strengthen school efficacy. To effectively manage the system, it is important to:

- ▶ Identify site-based monitoring staff members to maximize system credibility and buy-in, define their roles and responsibilities, and provide strong and ongoing training.
- ▶ Implement rigorous oversight using the monitoring materials and processes to ensure validity and reliability.

TASK 3

Show Continuous Improvement

The site-based management staff should engage stakeholders to leverage results that structure learning communities and to make the best use of resources to move from the individual to the department, school, or district level.

Continuous monitoring and evaluating of the system should lead to:

- ▶ Identification of structures to provide support at the broader level.
- ▶ Engagement of stakeholders to implement improvements to the evaluation system.

STAGE **3**

Conducting Collaborative Evaluations that Enhance Teaching Practice and Student Learning



The observation process should result in meaningful feedback that enhances teacher practice and student learning, improves the school's effectiveness, and promotes its values.

TASK 1

Define the Observation Process

At a minimum, observations should be conducted twice a year—once to provide the teacher an initial formative observation to facilitate feedback and support, and a second time to provide a final or summative evaluation and its associated results, along with a continuous growth plan. For example, the observation structure should establish

- ▶ ***Pre-observation requirements***, in which the teacher would provide multiple sources of evidence to provide contextual information prior to the observation (e.g., standards focus, knowledge of students, lesson selection, student work) so that the observer understands what the teacher plans to accomplish
- ▶ ***Observation requirements***, which establish the length of classroom observation, who conducts the observation, whether observations are announced or unannounced, the type of data collected, and an understanding of the data tool and its elements
- ▶ ***Post-observation requirements***, in which the observer opens the dialogue by providing feedback based on the observation and other sources of evidence; the teacher and observer review together the results of the observation and other measures (e.g., attendance and behavior data, standardized test results, and student learning artifacts), thus allowing the teacher to provide a response and written reflection; and they then collectively build a professional growth plan

Not all observations have to be in person; teachers could video classroom instruction to review with their peers or supervisors. Ongoing videos could be used as multiple measures of performance over time to support growth.

TASK 2

Communicate the Results of Observation Levels

All constituents must clearly understand the potential implications (e.g., interventions and career advancement) of a given level and under what circumstances they take effect. The observer must be equally adept in facilitating conversations about the positive and difficult topics when communicating the results. The summative evaluation determines the level and types of support designed for each teacher to enhance practice and student learning.

TASK 3

Specify the Evaluation Results

When determining the evaluation results, report elements could include

- ▶ Overall rating as determined by district—ineffective, adequate, proficient, exemplary
- ▶ Plan to strengthen effective practice—self-determined growth plan, 2-year growth plan, 1-year improvement plan

TASK 4

Specify the Consequences

Clearly define and communicate the actions or consequences related to rating decisions. Specify any connection to compensation, promotion opportunities, formal improvement plans, and termination, as well as the process to inform personnel decisions.

TASK 5

Specify the Appeals Process

Clearly establish an appeals process for teachers who consider their evaluation deficient, and include elements such as

- ▶ An appropriate body with decision-making authority
- ▶ Guidelines for reviewing ratings
- ▶ A protocol for providing results in a timely manner

STAGE **4**

Supporting Teachers



A key focus of this evaluation system is to create, promote, and sustain purposeful, collaborative teacher support that results in teacher growth and increased student learning. This support is critical to training new and developing teachers, retaining strong teachers, and enhancing teacher practice in all schools.

TASK 1

Specify Differentiated Support

When one is developing forward-focused teacher support, it is important to tailor support on the basis of the teacher's identified need. Support and professional development should be customized to particular contexts, student populations, developmental levels of students, teachers' needs, and school teaching and learning conditions (e.g., time, facilities, resources, school climate, safety). When support is built around those contexts, teaching practices will get more traction and momentum.

To guide the support efforts, one can use the rating levels from the observation to structure the requirements for types of support and subsequent observation. For example, for teachers with exemplary and proficient ratings, support could be self-directed with peer review using mentors or peer coaches, especially those who have earned National Board Certification. Conversely, for those teachers earning ineffective ratings, support might be defined by structured growth plans with supervisor oversight. Supervisors may assign instructional coaches to support teachers who have little time to seek solutions and research best practices to improve their own practice. An instructional coach could work hand in hand with teachers to help them improve their practice through continued support and follow-up.

Further, there is a great deal of evidence indicating that teachers learn best when they are members of learning communities—another factor that should be considered in structuring teacher professional development. Such communities can be school based and led by exemplary teachers and National Board Certified Teachers or can be built around learning networks of groups made up of teachers from different schools. Principals can provide teachers with continuous blocks

of time to be used in a variety of ways so teachers can teach other teachers the strategies that have been successful with their own students, technology to illustrate good teaching, school–university partnerships, and networked communities that support school-to-school partnerships, and they should build networks of teacher communities within schools and feeder patterns. In addition, support can include summer institutes that provide a set of social practices that can be used in the classroom. (Lieberman and Wood, 2003).

TASK 2

Monitor Effectiveness of Support

If one is to ensure that the support results in teacher growth and student learning, it is important to develop and implement a system to collect information such as teacher feedback on the effectiveness of programs and teachers' willingness to mentor others to expand and sustain the growth.

In a recent paper titled "Building Professional Development to Support New Student Assessment Systems", Learning Forward Executive Director Stephanie Hirsh writes that effective professional development "is only as effective as the outcomes it is intended to produce" and must be planned and managed in a collaborative manner with administrators and teacher leaders working together.

A coherent plan with measures for success puts professional development on the correct track by building an awareness sustained with long-term support that has coherence, duration, active learning, and collective participation and that is content focused. By incorporating teachers into the process, principals identify teachers' support needs in real time so that the intervention or information they require is available when they need it most.

School districts also must be willing to evaluate professional development strategies and to discard ineffective practices. If outside facilitators are used, they should be expected to explain what success will look like and then provide evidence of that success over time.

TASK 3

Highlight Use of Evaluation Results to Improve School Culture, Teacher Practice, and Student Outcomes

Leveraging observation results can help structure learning communities and can identify the types of mentors, peer coaches, and specialists that will best develop individual teacher practice and overall school efficacy. Outcomes of leveraging could include, but are not limited to,

- ▶ Developing the support systems to promote individual teacher growth and overall school improvement
- ▶ Bolstering and supporting effective practices and techniques
- ▶ Building and administering teacher survey instruments that help identify common frustrations and needs
- ▶ Identifying common issues to focus school or department-wide interventions
- ▶ Evaluating training and support efforts that are particularly effective or ineffective
- ▶ Identifying teacher leaders and using them as role models and mentors for effective practice
- ▶ Prioritizing areas of greatest need for interventions and support
- ▶ Comparing with similar schools and building networks to jointly address common needs
- ▶ Providing a network model in a setting where teachers can work together in a structured way on issues of instructional practice relevant to their work and can develop their understanding and skills around practices of improvement
- ▶ Structuring time for walk-arounds, practitioner rounds, peer support, or collaboration with key constituents

Outcomes

With the goal of improving the effectiveness of teacher practice and ultimately student learning, support should be available at the individual level as well as at a broader level—whether department, school, or district. Outcomes of having an effective teacher evaluation and support system may be

- ▶ Higher student performance
- ▶ Enhanced teacher practice
- ▶ Sustainable school growth
- ▶ Increased teacher retention
- ▶ Increased student engagement
- ▶ Collaborative learning communities
- ▶ Improved return on investment of professional development dollars
- ▶ Strengthened human capital in high-need schools

Resources

Many programs and states have developed components of the system as described in this comprehensive guide. Following is a short list of sources available for reference when developing a teacher evaluation framework:

- ▶ The **Performance Assessment for California Teachers (PACT)**, a system designed by a consortium of teacher preparation institutions, is an evaluation of performance expectations for pre-service teachers. The design of the common assessment was modeled, in part, after the portfolio assessments of the National Board.
- ▶ The **North Carolina Teacher Evaluation Process (NCTEP)** is aligned to the National Board's Five Core Propositions, thus accounting for knowledge of content, knowledge and management of students and for reflective practice.
- ▶ Massachusetts has revised its **Massachusetts Triangulated Standards-Based Evaluation Framework** to focus on professional practice that provides feedback and support to teachers. When teachers were surveyed in 2010, only one-third indicated that the evaluation system identified specific professional development opportunities. To guide deeper, more substantive evaluation, the state has developed new standards and indicators for teacher practice that are grounded in the National Board's Five Core Propositions. The new system continues to include "observation of practice and examination of artifacts, but incorporates two new elements: validation of the initial judgment by examining measures of student learning and outcomes and assessing evidence of professional contributions" (Skinner, 2010).

- ▶ The **Continuous Improvement Model for Teacher Development and Evaluation**, developed by the American Federation of Teachers, proposes a way to systematically improve teacher practice and increase student learning through regular, rigorous reviews by trained evaluators. Based on professional teaching standards, best practices, and student achievement, effective evaluations can improve public education by growing promising teachers, enabling good teachers to become great, and identifying teachers who shouldn't be in the classroom at all.
- ▶ Based on its vision of great public schools for every child, the National Education Association (NEA) has identified the core purposes and values of a comprehensive teacher growth and development system to meet the demands of the 21st century. **Teacher Assessment and Evaluation** provides a framework designed to transform education from a reward and punish system to one that supports effective teaching and student learning. The framework describes guiding principles and lays out a process for developing a comprehensive and robust teacher evaluation system. In 2011, NEA also set forth a policy statement describing the types of teacher evaluation and accountability systems necessary to ensure a high quality public education for every student.

Conclusion

All teachers—and all professionals—seek meaningful evaluation, feedback, and support. When concluding a key overarching lesson learned during the 2011 Summit, *Improving Teaching Quality Around the World* reported, “The quality of an education system cannot exceed the quality of its teachers—but neither will the quality of teaching exceed the quality of the systems in place to recruit, train, develop, and advance teachers.”

The National Board invites policymakers and officials at the state and district levels to give careful consideration to this framework in determining how teachers will be evaluated and supported. This document is intended to bring clarity to the field. Although the details of an evaluation system must be worked out locally, the stages provide the framework to effectively guide the process. If we get this right as a nation, we will strengthen, retain, and expand our teacher workforce and will thereby improve student performance.

The need to effectively evaluate teachers is urgent, though it is just as important that we get it right.

Bibliography

- American Federation of Teachers. *Continuous Improvement Model for Teacher Development and Evaluation*. Washington, DC: Author, 2010.
- Asia Society Partnership for Global Learning. *Improving Teacher Quality Around the World: The International Summit on the Teaching Profession*. New York: Asia Society, 2011.
- Baker, E., et al. *Problems with the Use of Student Test Scores to Evaluate Teachers*. Washington, DC: Economic Policy Institute, 2010.
- Darling-Hammond, L. *Evaluating Teacher Effectiveness: How Teacher Performance Assessments Can Measure and Improve Teaching*. Washington, DC: Center for American Progress, 2010.
- Elmore, R. F. "Professional Networks and School Improvements." *The School Administrator* 64, no. 4 (April 2007): 20–24.
- Grossman, T. *Building a High-Quality Education Workforce: A Governor's Guide to Human Capital Development*. Washington, DC: NGA Center for Best Practices, 2009.
- Harris, D. N. *Value-Added Measures in Education: What Every Educator Needs to Know*. Cambridge, MA: Harvard Education Press, 2011.
- Hirsh, S. *Building Professional Development to Support New Student Assessment Systems*. Oxford, OH: Learning Forward, 2010.
- Ingvarson, L., and J. Hattie. *Assessing Teachers for Professional Certification: The First Decade of the National Board for Professional Teaching Standards*. Amsterdam: JAI/Elsevier Science, 2008.
- Lieberman, A., and D. H. Pointer Mace. *Teacher Learning: The Key to Educational Reform*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching, 2008.
- Lieberman, A., and D. R. Wood. *Inside the National Writing Project: Connecting Network Learning to Classroom Teaching*. New York: Teachers College Press, 2003.
- Mertler, C. A. "Designing Scoring Rubrics for Your Classroom." *Practical Assessment, Research & Evaluation* 7, no. 25 (2001) Retrieved June 6, 2011, from <http://PAREonline.net/getvn.asp?v=7&n=25>.
- Moskal, B. M. "Scoring Rubrics: What, When and How?" *Practical Assessment, Research & Evaluation* 7, no. 3 (2000). Retrieved June 6, 2011 from <http://PAREonline.net/getvn.asp?v=7&n=3>.

National Board for Professional Teaching Standards. *Student Learning, Student Achievement: How Do Teachers Measure Up?* Arlington, VA: National Board for Professional Teaching Standards, 2011.

National Board for Professional Teaching Standards. *What Teachers Should Know and Be Able to Do*. Washington, DC: Author, 1994.

National Board Resource Center. *A Quality Teacher in Every Classroom: Creating a Teacher Evaluation System That Works for California*. Stanford, CA: Author, 2010.

National Education Association. *Teacher Assessment and Evaluation: The National Education Association's Framework for Transforming Education Systems to Support Effective Teaching and Improve Student Learning*. Washington, DC: Author, 2010.

National Education Association. *Policy Statement on Teacher Evaluation and Accountability*. 2011. [Online]. Available from <http://www.nea.org/grants/46326.htm>. 30 September 2011.

Skinner, K. J. *Reinventing Educator Evaluation: Connecting Professional Practice with Student Learning*. Boston: Massachusetts Teachers Association, 2010.

Stiggins, R. J. *Student-Involved Classroom Assessment*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2001.

Weisberg, D., S. Sexton, J. Mulherm, and D. Keeling. *The Widget Effect: Our National Failure to Act on Differences in Teacher Effectiveness*. Brooklyn, NY: The New Teacher Project, 2009.

©2011 National Board for Professional Teaching Standards. All rights reserved, NBPTS, NBCT, National Board for Professional Teaching Standards, National Board Certified Teacher, National Board Certification, *Take One!* and Accomplished Teacher are registered trademarks or service marks of the National Board for Professional Teaching Standards. Other marks are trademarks or registered trademarks of their respective organizations.

The contents of this publication were developed in whole or in part under a grant from the U.S. Department of Education. However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the federal government.



National Board for Professional Teaching Standards
1525 Wilson Boulevard, Suite 500 · Arlington, VA 22209
www.growgreatschools.org · 1-888-780-7805