

TEACHER EVALUATION

2.0

Annual process
Clear, rigorous expectations
Multiple measures
Multiple ratings
Regular feedback
Significance

Everyone agrees that teacher evaluations are broken. So how can we fix them? This guide proposes **six design standards** that any rigorous and fair evaluation system should meet. It offers states and school districts a blueprint for better evaluations that can help every teacher thrive in the classroom—and give every student the best chance at success.

The next several years represent a golden opportunity to create better systems that meet the needs of schools and the professionals who work in them: **Teacher Evaluation 2.0.**

TEACHERS MATTER

Years of research have proven that nothing schools can do for their students matters more than giving them effective teachers. A few years with effective teachers can put even the most disadvantaged students on the path to college. A few years with ineffective teachers can deal students an academic blow from which they may never recover.*

“The effect of increases in teacher quality swamps the impact of any other educational investment, such as reductions in class size.” *Goldhaber, 2009*

“More can be done to improve education by improving the effectiveness of teachers than by any other single factor.”
Wright, Horn and Sanders, 1997

“Having a top-quartile teacher rather than a bottom-quartile teacher four years in a row could be enough to close the black-white test score gap.” *Gordon, Kane and Staiger, 2006*

“Having a high-quality teacher throughout elementary school can substantially offset or even eliminate the disadvantage of low socio-economic background.”
Rivkin, Hanushek and Kain, 2002

Research has also shown that the best predictor of a teacher’s effectiveness is his or her past success in the classroom. Most other factors pale in comparison, including a teacher’s preparation route, advanced degrees, and even experience level (after the first few years). The lesson is clear: to ensure that every child learns from the most effective teachers possible, schools must be able to gauge their teachers’ performance fairly and accurately.

*Jordan, Mendro, and Weerasinghe, *The Effects of Teachers on Longitudinal Student Achievement*, 1997

TEACHER EVALUATION 2.0

Nearly everyone agrees that great teachers are critical to student success—and that our schools have not done nearly enough to evaluate teachers accurately and use this information to improve educational quality.

Increasingly, school districts, states and teachers' unions are advancing evaluation reform through legislation and by negotiating changes to collective bargaining agreements. This has compelled education leaders and policymakers to grapple with difficult issues that have received only lip service in the past: How can we help all teachers reach their full potential in the classroom? How can we ensure that teachers love their jobs, so that the best teachers want to keep teaching? How can we address consistently ineffective teaching fairly but decisively?

We cannot address any of these issues without better teacher evaluation systems.

Evaluations should provide all teachers with regular feedback that helps them grow as professionals, no matter how long they have been in the classroom. Evaluations should give schools the information they need to build the strongest possible instructional teams, and help districts hold school leaders accountable for supporting each teacher's development. Most importantly, they should focus everyone in a school system, from teachers to the superintendent, on what matters most: keeping every student on track to graduate from high school ready for success in college or a career.

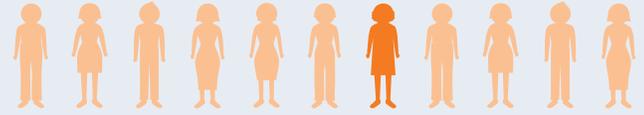
Evaluations should do all of these things, but in most cases, they don't even come close. Instead, they are typically perfunctory compliance exercises that rate all teachers "good" or "great" and yield little useful information. As Secretary of Education Arne Duncan noted in a summer 2010 speech, "our system of teacher evaluation... frustrates teachers who feel that their good work goes unrecognized and ignores other teachers who would benefit from additional support."

The next several years represent a golden opportunity to create better systems that meet the needs of schools and the professionals who work in them: Teacher Evaluation 2.0.

Inspired by the federal Race to the Top competition, states and districts across the country have begun to revamp outdated evaluation systems. Teachers' unions have shown a willingness to become partners in this work; the American Federation of Teachers, for example, recently awarded grants to local chapters that are helping to design new evaluation systems.

The crucial question now facing education leaders is, "How?" How can they avoid the pitfalls of past evaluation systems? How can they create evaluations that become useful tools for teachers and school leaders, and that help push students to new heights? What can they learn from the districts and states that are making real progress?

This guide is intended to address these critical questions. We hope to provide a blueprint for rigorous, fair and credible teacher evaluation systems centered on student outcomes.



PROBLEMS WITH CURRENT EVALUATION SYSTEMS

As we showed in our 2009 report, *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*, most teacher evaluation systems suffer from a slew of design flaws.

Infrequent: Many teachers—especially more experienced teachers—aren't evaluated every year. These teachers might go years between receiving any meaningful feedback on their performance.

Unfocused: A teacher's most important responsibility is to help students learn, yet student academic progress rarely factors directly into evaluations. Instead, teachers are often evaluated based on superficial judgments about behaviors and practices that may not have any impact on student learning—like the presentation of their bulletin boards.

Undifferentiated: In many school districts, teachers can earn only two possible ratings: "satisfactory" or "unsatisfactory." This pass/fail system makes it impossible to distinguish great teaching from good, good from fair, and fair from poor. To make matters worse, nearly all teachers—99 percent in many districts—earn the "satisfactory" rating. Even in districts where evaluations include more than two possible ratings, most teachers earn top marks.

Unhelpful: In many of the districts we studied, teachers overwhelmingly reported that evaluations don't give them useful feedback on their performance in the classroom.

Inconsequential: The results of evaluations are rarely used to make important decisions about development, compensation, tenure or promotion. In fact, most of the school districts we studied considered teachers' performance only when it came time to dismiss them.

Taken together, these shortcomings reflect and reinforce a pervasive but deeply flawed belief that all teachers are essentially the same—interchangeable parts rather than individual professionals.

For more information, visit www.tntp.org/widget.



GUIDING PRINCIPLES

A meaningful teacher evaluation system should reflect a set of core convictions about good instruction. Unfortunately, most evaluations communicate a devastating message—that all teachers are about the same, and that the primary purpose of evaluation is to identify and remove a tiny number of teachers who are judged grossly incompetent. The typical evaluation form suggests that good teaching consists of performing a mundane set of routines that are largely unrelated to student engagement or learning. The standards we propose for Teacher Evaluation 2.0 are founded on a far different set of core principles about the power of great teachers and the critical role evaluations play in developing them:

All children can master academically rigorous material, regardless of their socioeconomic status. A teacher who believes his or her students cannot meet ambitious expectations is not the right fit for that classroom. Great teachers across the country prove every day that students can consistently succeed in spite of enormous challenges outside the classroom. Furthermore, it is possible to set reasonable targets for the amount of academic progress each student should be able to make in a year, taking into account the student's academic history.

A teacher's primary professional responsibility is to ensure that students learn. Therefore, measures of student learning should play a predominant role in teacher evaluations. This does not mean that teacher evaluations should be based solely on the results of standardized tests, or based on the results of any single assessment. But it *does* mean that teachers should be accountable for helping students make measurable progress against ambitious learning standards.

Teachers contribute to student learning in ways that can largely be observed and measured. Through focused, rigorous observation of classroom practice, examination of student work, and analysis of students' performance on high-quality assessments, it is possible to accurately distinguish effective teaching from ineffective teaching. Great teachers vary widely in their instructional style and approach, but they all share a powerful ability to nurture student academic growth.

Evaluation results should form the foundation of teacher development. Although there must be meaningful consequences for consistently poor performance, the primary purpose of evaluations should not be punitive. Good evaluations identify excellent teachers and help teachers of all skill levels understand how they can improve; they encourage a school culture that prizes excellence and continual growth. With better teacher evaluations in place, school districts can also do a better job holding *school leaders* accountable for doing their most important job: helping teachers reach their peak. Removing persistently underperforming teachers is a necessary but insufficient step to building a thriving teacher workforce.

Evaluations should play a major role in important employment decisions. If we want good teaching in every classroom, good teaching must be valued. District leaders should factor teachers' effectiveness—as measured by evaluations—into decisions about hiring, pay increases, promotions, tenure and retention. The goal is not to increase teachers' level of effort or penalize struggling teachers, but to make teaching a fulfilling career and a profession that talented people aspire to enter and master.

No evaluation system can be perfect—in teaching or in any other profession. But we can develop systems that are dramatically better than current ones, and that teachers and school leaders believe are fair and accurate. Once we do, we should use them and improve on them.

DESIGN STANDARDS

With these guiding principles and the flaws of current evaluation systems in mind, we propose six design standards that any teacher evaluation system must meet in order to be effective. These six standards are interdependent; each is critical to ensuring that evaluations meet the needs of teachers, school leaders and students. Each standard is described in detail in the following pages, along with real-life examples and potential pitfalls.

- 1 ANNUAL PROCESS**
All teachers should be evaluated at least annually.
- 2 CLEAR, RIGOROUS EXPECTATIONS**
Evaluations should be based on clear standards of instructional excellence that prioritize student learning.
- 3 MULTIPLE MEASURES**
Evaluations should consider multiple measures of performance, primarily the teacher's impact on student academic growth.
- 4 MULTIPLE RATINGS**
Evaluations should employ four to five rating levels to describe differences in teacher effectiveness.
- 5 REGULAR FEEDBACK**
Evaluations should encourage frequent observations and constructive critical feedback.
- 6 SIGNIFICANCE**
Evaluation outcomes must matter; evaluation data should be a major factor in key employment decisions about teachers.

1 | ANNUAL PROCESS

School leaders should evaluate every teacher at least once a year. Annual evaluation is the only way to ensure that all teachers—regardless of their ability level or years of experience—get the ongoing feedback on their performance that all professionals deserve. This approach recognizes that a teacher’s effectiveness and developmental needs may change over time, and it sends a message to school leaders that they are accountable for helping *all* their teachers grow as professionals. The ratings from annual evaluations will also allow schools to make important employment decisions based on up-to-date information.

// PITFALL TO AVOID // *Attempting to save resources by evaluating veteran teachers with no performance issues less frequently.* In some districts, veteran teachers go as many as five years between full evaluations. This policy is based on the flawed assumption that professional growth happens only at the very beginning of a teacher’s career—which becomes a self-fulfilling prophecy without regular feedback. While teachers deserve special attention during their first years in the classroom, when they grow the most, all teachers develop over time and deserve recognition for their successes and help with their challenges. If we care about what happens in the classroom, we should care more often than every five years.

2 | CLEAR, RIGOROUS EXPECTATIONS

Teachers should be evaluated against clear, rigorous performance expectations based primarily on evidence of student learning (as opposed to teacher behaviors or routines). Expectations should reflect excellence in the classroom, not minimally acceptable performance. They should also be precisely worded and leave little room for inference (e.g., be built around observable evidence that students are actively engaged in the lesson, not whether the lesson plan appears to be engaging on paper), to ensure that all teachers and instructional managers interpret them in the same way.

Districts should create clear, concise tools to help instructional managers consistently evaluate teachers against these expectations. For example, observation rubrics should be aligned closely to performance expectations and should address specific, observable student behaviors that school leaders can document while watching a lesson (e.g., approximate percentage of students who raise their hands when the teacher poses a question).

Expectations should also be refined regularly, especially during the first few years that they are in use. Any elements that are not directly related to student achievement should show a strong correlation to student achievement over time or be discarded.

// PITFALL TO AVOID // *Setting the bar for “ineffective” too low.* In defining “ineffective” teaching, some districts include only blatantly negligent, harmful instruction that reflects a near-total lack of effort or skill—a depth to which almost no well-intentioned teacher drops. In reality, ineffective instruction can take many forms; it is reflected more in the responses of students than in the behaviors of teachers. A lack of student engagement and academic progress indicates ineffective teaching, regardless of what the teacher is doing.

GREAT EXPECTATIONS: ACHIEVEMENT FIRST CHARTER SCHOOL NETWORK

The teacher evaluation process at Achievement First (AF) is a model of rigorous, low-inference performance expectations coupled with regular feedback for teachers. Expectations at AF—called “The Essentials of Effective Instruction”—help instructional managers make connections between students’ progress and the specific actions of their teacher that may be contributing to that progress. Additionally, the low ratio of teachers to instructional managers in AF schools (usually no higher than 8:1) allows teachers to receive regular, intensive coaching on development areas. Instructional managers observe all their teachers and provide feedback at least once every two weeks, on average, and development plans are updated every six weeks to address teachers’ most critical needs. The frequency and quality of support that teachers receive at AF is possible because of career paths that allow the most effective AF teachers to become “home grown” instructional managers.

Although AF’s leaders describe this evaluation process as a work in progress that is constantly improving, its early success is proof that it is possible for schools to provide every teacher with frequent, individualized support to meet rigorous expectations.

Expectations should reflect excellence in the classroom, not minimally acceptable performance.

EXPECTATIONS EXPLAINED

Poorly constructed expectation:

Teacher makes a thoughtful and accurate assessment of the level of student understanding throughout the lesson. (General, teacher-centered, open to interpretation)

Clear and rigorous expectation:

Student work during the course of the lesson (e.g., “Do Nows,” checks for understanding, guided and independent practice, exit slips) shows that nearly all students at all skill levels mastered the lesson objectives. (Specific, student-centered, little room for inference)

EXPECTATIONS AND EXPERIENCE

While it is essential to hold high expectations for all teachers from their earliest days in the classroom, it is unrealistic to expect even talented novice teachers to meet the same expectations as more practiced educators. For this reason, expectations should increase steadily during a teacher’s first three years in the classroom—the time when the greatest amount of improvement typically occurs. This will allow for an easy assessment of an early-career teacher’s development trajectory, so that school leaders can determine whether the teacher is on track to meet the district’s ultimate expectations within a reasonable amount of time. For example, a district might renew a first-year teacher who is rated “Needs Improvement” because she is improving at a reasonable rate and is just beginning her career. If that teacher was still rated “Needs Improvement” when she came up for tenure, however, she might be subject to dismissal.

SUMMATIVE RATING & RENEWAL GUIDANCE

EXPERIENCE (in years)	INEFFECTIVE	NEEDS IMPROVEMENT	EFFECTIVE	HIGHLY EFFECTIVE
1	Renew if improving rapidly	Renew if improving	Renew	Renew
2	Do not renew	Renew if improving	Renew	Renew
3	Deny tenure	Extend probationary period	Grant tenure	Grant tenure
4+	Dismiss or counsel out	Retain if the teacher was “effective” last year; otherwise dismiss or counsel out	Retain	Retain and reward

3 | MULTIPLE MEASURES

No single data point can paint a complete picture of a teacher’s performance, so evaluation systems should use multiple measures to determine whether teachers have met performance expectations. Whenever possible, these should include objective measures of student academic growth, such as value-added models that connect students’ progress on standardized assessments to individual teachers while controlling for important factors such as students’ academic history. Other possible measures include performance on district-wide or teacher-generated assessments, and classroom observations centered on evidence of student learning.

Each measure should have a specific weight, so that teachers and instructional managers understand how each component will factor into the final evaluation rating. The most weight should be afforded to the most accurate measures of student progress, which will often be the objective measures.

// PITFALL TO AVOID // **Treating all data sources as equally valid.** Some assessments of student progress are more reliable than others. For example, a writing prompt that is administered district-wide and is designed to probe several key learning standards generally yields more useful data for evaluations than a teacher-designed take-home essay. Although the teacher-designed essay may offer useful information about student learning that could be used in the evaluation process, district-wide assessments typically ensure more consistent, rigorous standards and allow for better comparisons between teachers. Multiple measures should not mean “anything goes.”

WEIGHTING MULTIPLE MEASURES OF TEACHER PERFORMANCE

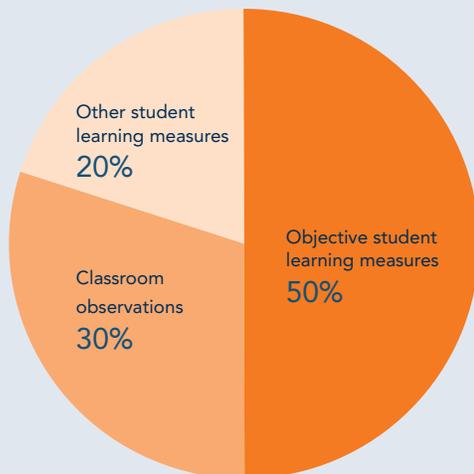
The illustrations below show how various potential measures of teacher performance might be weighted when teacher value-added data are and are not available. These illustrations are based on three primary measures of teacher performance:

Objective student learning measures, such as value-added results.

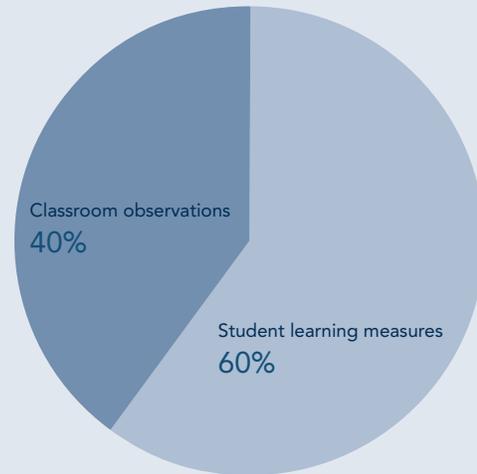
Classroom observations, including analysis of instructional practice, lesson planning, professionalism and student engagement, among other possible factors.

Other student learning measures, such as progress toward Individual Education Plan (IEP) goals, district-wide or teacher-generated assessments, and end-of-course tests.

SAMPLE EVALUATION WEIGHTS



VALUE-ADDED DATA AVAILABLE



NO VALUE-ADDED DATA AVAILABLE

These weights are illustrative and will vary based on the specific measures that become part of the evaluation system. The most accurate measures of student progress should carry the greatest weight to ensure that a teacher’s overall rating is as accurate as possible. In deciding on weights, it is important to simulate all possible outcomes on the different measures to ensure that each combination leads to a sensible overall rating. This analysis may produce surprising results and lead to refinements in the weighting of certain measures. In some cases, a summative evaluation rating matrix (see next page) may be a better alternative for determining weights.

4 | MULTIPLE RATINGS

Each teacher should earn one of four or five summative ratings at the end of each school year: for example, “highly effective,” “effective,” “needs improvement” or “ineffective.”

This number of categories is large enough to give teachers a clear picture of their current performance, but small enough to allow for clear, consistent distinctions between each level and meaningful differentiation of teacher performance within schools and across the district. For example, instructional managers will need to think carefully about whether a moderately performing teacher falls into the “effective” or “needs improvement” category, which will ensure that the teacher receives support tailored to her specific needs.

// PITFALL TO AVOID // **Ambiguity about the meaning of each rating.** Rating scales should include at least two levels at or above expectations and two levels below expectations, and there should be no ambiguity about which levels represent meeting expectations. Teachers deserve clear information about whether they are meeting the district’s overall expectations, and school leaders need the same clarity in order to use evaluations wisely in making important employment decisions.

MULTIPLE MEASURES, MULTIPLE RATINGS— AND MULTIPLE COLLABORATORS: NEW HAVEN PUBLIC SCHOOLS’ TEACHER EVALUATION SYSTEM

As part of a new collective bargaining agreement, New Haven Public Schools (NHPS) worked closely with the local teachers’ union in the winter of 2009-10 to design a new evaluation system centered on evidence of student learning. Under the new system, all teachers receive ratings in two overarching categories: “student learning growth,” which reflects student progress on standardized and teacher-generated assessments, and “instructional practice/professional values,” which reflects regular observations by administrators on factors such as lesson planning and preparation, classroom practice, use of data, professionalism, and high expectations for students.

In coordination with the union, NHPS developed a matrix that clearly illustrates how the ratings in each category will combine to produce a summative rating on a scale from “needs improvement” to “exemplary.” To help ensure that the new system is implemented consistently, impartial observers must review administrators’ judgments about low- and high-performing teachers. The district and the union agreed that teachers who consistently earn low ratings will be dismissed, while teachers who consistently earn high ratings will be eligible for promotions or other recognition. The new evaluation system is being implemented for the first time during the 2010-11 school year. NHPS is proof that meaningful evaluation systems can emerge from focused collaboration between school districts and teachers’ unions.

NEW HAVEN PUBLIC SCHOOLS SUMMATIVE EVALUATION RATING MATRIX

STUDENT LEARNING GROWTH

		Needs Improvement (NI-1)	Developing (Dv-2)	Effective (Ef-3)	Strong (St-4)	Exemplary (Ex-5)
INSTRUCTIONAL PRACTICE AND PROFESSIONAL VALUES	Needs Improvement (NI-1)	NI-1	NI-1	Dv-2	Ef-3*	Ef-3*
	Developing (Dv-2)	NI-1	Dv-2	Dv-2	Ef-3	St-4*
	Effective (Ef-3)	NI-1	Dv-2	Ef-3	St-4	Ex-5
	Strong (St-4)	Dv-2*	Ef-3	St-4	St-4	Ex-5
	Exemplary (Ex-5)	Ef-3*	Ef-3*	St-4	Ex-5	Ex-5

*Ratings with this degree of mismatch are subject to review by the NHPS central office.

Note: TNTP assisted in the design of this evaluation system.

5 | REGULAR FEEDBACK

An evaluation system should not be limited to a single rating assigned at the end of the year.

Instead, instructional managers should strive to cultivate a performance-focused culture by observing their teachers frequently. They should also have regular conversations with their teachers to discuss overall classroom performance and student progress; professional goals and developmental needs; and the support school leaders will provide to meet those needs. Teachers and instructional managers should come away from these conversations with a shared understanding of what the teacher needs to focus on in the short term and how the instructional manager will help. If teachers are surprised by their summative evaluation rating, something is wrong with the evaluation process.

// PITFALL TO AVOID // Compliance-driven feedback. Feedback is useless if instructional managers and teachers view development conversations as chores instead of opportunities to talk openly and constructively about instruction. Districts should hold instructional managers accountable for the *quality* of the feedback and support teachers receive, not just the quantity (as we discuss in the next design standard). Additionally, an evaluation should not become void if an administrator misses even one deadline in the process, as is the case in too many districts today. Compliance with the major elements and spirit of the process should be sufficient, and administrators who consistently fall short of this goal should be held accountable in their own evaluations.

6 | SIGNIFICANCE

An evaluation process must have meaningful implications, both positive and negative, in order to earn sustained support from teachers and school leaders and to contribute to the systematic improvement of the teacher workforce.

It should produce information that districts can easily factor into important decisions about teacher tenure, compensation, development, hiring, promotion and dismissal. This means that the results of evaluations must be accurate, clear and easy to interpret.

Some policymakers may support using evaluations only to reward excellent teachers, and not for more difficult decisions like layoffs. But if teacher performance matters at all, it should matter for any significant decision that affects the quality of instruction students receive. As schools seek to build and sustain strong instructional teams, a teacher's track record of success in the classroom should be paramount, not off-limits.

Accountability for evaluation outcomes should not rest on the shoulders of teachers alone; the ability to identify, develop and keep talented teachers is arguably the most important priority of any school leader. Therefore, instructional managers should be held accountable not just for evaluating teachers accurately, but for acting on the results and helping teachers improve over time. Likewise, district and state education leaders should be similarly responsible for ensuring that instructional managers receive effective oversight and the training and support they need to evaluate teachers fairly, consistently and accurately.

// PITFALL TO AVOID // Blind allegiance to evaluation results.

While evaluations should always be a significant part of employment decisions, professional judgment must play a role, too. For example, the highest performing math teacher should not automatically be selected as a math coach, regardless of his or her other qualities. Likewise, a teacher should not be summarily dismissed after a single negative classroom observation. Each incremental change in performance does not require an instant and significant reward or penalty.

EVALUATION OUTCOMES WITH SIGNIFICANCE: WASHINGTON, DC PUBLIC SCHOOLS' IMPACT SYSTEM

In just two years, the District of Columbia Public Schools (DCPS) designed a new evaluation system, piloted it, implemented it system-wide and used the results to make key employment decisions. The new system, IMPACT, evaluates teachers based on their students' growth on standardized assessments and on frequent observations by both administrators and impartial master teachers who are not managed by principals. Observations are tied to a clear, detailed and rigorous rubric that focuses on evidence of student learning. Evaluators meet regularly with teachers to give feedback.

In IMPACT's first full year of implementation, schools differentiated their teachers' performance much more than in previous years, and DCPS took action based on the results. Sixteen percent of teachers who were rated "highly effective" will earn bonuses under the district's new compensation system. The 5 percent of teachers who earned an "ineffective" rating were dismissed. Another 17 percent were rated "minimally effective" and may be dismissed in a year if they fail to improve after receiving intensive support.

Although IMPACT was not implemented easily or without controversy, DCPS proved that it is possible for a large district to move from design to implementation—complete with rewards and consequences—in just two years.

IMPLEMENTATION

The success of any evaluation system—no matter how solid its design—ultimately depends on how well it is implemented.



For example, school leaders and their managers will need ongoing training and support on the technical aspects of the system and other performance management issues, like having constructive conversations with teachers about performance concerns. Teachers will need clear information about how the system works and how they can suggest improvements. This will likely require directing more resources and personnel toward teacher evaluations and relieving administrators of less critical responsibilities.

Furthermore, even the most elegantly designed system will need to be improved over time. That's why every district should establish specific metrics to track whether its evaluation system is functioning correctly and yielding its desired outcomes—both at individual schools and district-wide. Based on these metrics, district leaders should make any necessary adjustments to the design or implementation of the evaluation system every year. Questions to consider in developing these metrics include:

Are school leaders evaluating teachers accurately?

The distribution of summative evaluation ratings should roughly mirror patterns of student academic growth.

Are teachers generally improving their performance over time? Teachers—especially novice teachers—should improve to become “effective” or “highly effective.” Since holding school leaders accountable for this alone may encourage rating inflation, districts should validate ratings by using external evaluators or comparing ratings to objective evidence that a teacher is or is not improving over time (e.g., change in value-added percentile).

Are schools retaining consistently top-performing teachers at higher rates than consistently low-performing teachers? Districts should set specific goals for retaining teachers who earn top ratings for two or more consecutive years, with special emphasis on those who teach high-need students. School leaders should also be expected to make a compelling argument for every teacher they retain who earns consistently low ratings.

Are teachers receiving useful feedback based on clear expectations?

Districts should survey teachers regularly to ask whether they feel their school sets clear expectations for them and helps them meet those expectations. School leaders whose teachers consistently express dissatisfaction should be subject to additional scrutiny of their evaluation practices.

Do teachers believe they are being evaluated fairly?

Districts should survey teachers regularly to ask whether they are confident in the fairness and consistency of the evaluation process. As above, districts should investigate schools where larger percentages of teachers express concern.

Are school leaders getting the support they need to conduct accurate evaluations? Districts should survey school leaders regularly to ask whether they have the training, time and resources they need to implement the evaluation system well. District leaders, human resources staff and other support personnel should be held accountable when school leaders say they are not getting what they need.

The design standards and outcome metrics proposed here will set states and districts up for success as they begin to create Teacher Evaluation 2.0. Getting evaluations right is hard work, but the payoff is well worth the effort. Teachers will finally receive the feedback, support and recognition they deserve as professionals. School leaders will finally have the information and encouragement they need to make informed decisions as they build their instructional teams. Most importantly, all students will finally have access to the most important resource a school can provide: an effective teacher.



About The New Teacher Project

The New Teacher Project (TNTP) works to end the injustice of educational inequality by providing excellent teachers to the students who need them most and by advancing policies and practices that ensure effective teaching in every classroom. A national nonprofit organization founded by teachers, TNTP is driven by the knowledge that although great teachers are the best solution to educational inequality, the nation's education systems do not sufficiently prioritize the goal of effective teachers for all. In response, TNTP develops customized programs and policy interventions that enable education leaders to find, develop and keep great teachers and achieve reforms that promote effective teaching.

Since its inception in 1997, TNTP has recruited or trained approximately 37,000 teachers—mainly through its highly selective Teaching Fellows™ programs—benefiting an estimated 5.9 million students. TNTP has also released a series of acclaimed studies of the policies and practices that affect the quality of the nation's teacher workforce, most recently including *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness* (2009). Today TNTP is active in more than 40 cities, including Baltimore, Chicago, Denver, New Orleans, New York, and Oakland, among others.

www.tntp.org • info@tntp.org.