

MEASURING TEACHER EFFECTIVENESS

A LOOK “UNDER THE HOOD” OF
TEACHER EVALUATION IN 10 SITES

A CONNCAN, 50CAN, AND PUBLIC IMPACT REPORT



MEASURING TEACHER EFFECTIVENESS

A LOOK “UNDER THE HOOD” OF
TEACHER EVALUATION IN 10 SITES

BY DANIELA DOYLE AND JIYE GRACE HAN,
PUBLIC IMPACT

A CONNCAN, 50CAN, AND PUBLIC IMPACT REPORT

ACKNOWLEDGMENTS

ConnCAN, 50CAN, and Public Impact are grateful to the many partners and collaborators from the sites featured in this report: Sarah Coon of Achievement First; Madeleine Bayard of the Rodel Foundation; Diane Donohue of the Delaware Department of Education; David Steele of Hillsborough County Public Schools; Ellen Hur of The New Teacher Project; Michele Sherban-Kline of New Haven Public Schools; Sam Franklin, Abigail Loughrey, and Erika Gentry of Pittsburgh Public Schools; Brent Maddin, Thackston Lundy, Norman Atkins, and Kaycee Salmacia of the Relay Graduate School of Education; Melissa Cabral of the Rhode Island Department of Education; Emily Barton and Sara Heyburn of the Tennessee Department of Education; Kim Levensgood of District of Columbia Public Schools; Greg Sommers of the Bill & Melinda Gates Foundation; and Sandi Jacobs of the National Council on Teacher Quality. We also would like to thank Public Impact, including Bryan Hassel who advised on the report and Sharon Kebschull Barrett for copy-editing. Additional thanks go to Jennifer Alexander and Alex Spurrier from ConnCAN and Gretchen Guffy from 50CAN for their contributions to this report. This report was made possible by the generous support of the H. A. Vance Foundation.

PLEASE CITE THIS REPORT AS:

Doyle, D., & Han, J.G. (2012). *Measuring teacher effectiveness: A look “under the hood” of teacher evaluation in 10 sites*. New York: 50CAN; New Haven, CT: ConnCAN; and Chapel Hill, NC: Public Impact. Retrieved from <http://www.conncan.org/learn/research/teachers/measuring-teacher-effectiveness>

©2012 50CAN, ConnCAN, and Public Impact

The authors encourage the free use, reproduction, and distribution of this paper for noncommercial use. **We require attribution for all use.** For more information and instructions on the commercial use of our materials, please contact us at www.publicimpact.com.

TO ORDER ADDITIONAL COPIES:

Please contact ConnCAN at info@conncan.org or 203 772 4017

ConnCAN
85 Willow Street
New Haven, CT 06511

www.conncan.org
[@conncan](https://twitter.com/conncan)

CONTENTS

FOREWORD	1
INTRODUCTORY BRIEF	3
CROSS-SITE ANALYSIS	17
SITE PROFILES	25
DELAWARE	25
RHODE ISLAND	36
TENNESSEE	46
HILLSBOROUGH COUNTY, FL	53
HOUSTON, TX	60
NEW HAVEN, CT	69
PITTSBURGH, PA	78
WASHINGTON, D.C.	86
ACHIEVEMENT FIRST (CMO)	96
RELAY GRADUATE SCHOOL OF EDUCATION	105
MORE INFORMATION	115



3

FOREWORD

PATRICK RICCARDS, CONNCAN
MARC PORTER MAGEE, 50CAN

2

1

INCH

0

Research and experience make this clear: Great teachers change lives. They inspire and motivate students, and set them on a path for future success. By contrast, just one underperforming teacher can have a lasting negative impact on a student.

Given this reality, significant time and attention has rightly been focused on ensuring that all children have outstanding teachers at the front of their classrooms. This includes improving how teacher performance is evaluated and using evaluations to provide training, support, and recognition for outstanding performance. It also means ensuring swift and fair dismissal for the small number of teachers who do not improve after receiving support.

A number of publications have reported on state and district policies in this area and offered useful guidance for policy development. Little has been reported to date, however, about what happens once these policies are in place and the difficult work of design and implementation begins. Leaders, advocates, educators, and others must answer complex technical questions about how evaluation systems will work in practice. For example: “How should we evaluate a teacher in a non-tested grade or subject?” or “What mechanisms exist to ensure validity?”

Answering these questions can be a daunting task, but this work does not require reinventing the wheel. Those considering embracing rigorous teacher evaluations can learn from the states, districts, and education organizations that are already engaged in this work. That’s why we developed this report. In partnership with Public Impact, and with generous support from the H.A. Vance Foundation, Measuring Teacher Effectiveness looks “under the hood” at the evaluation systems being implemented by 10 leading sites, including states, districts, a charter management organization, and a graduate school of education.

None of these systems is perfect, but they do show us possible paths forward. Each site continually—and rightly—refines and improves its evaluation system. As other states and districts take on this work, we hope this information will help education leaders develop their own paths forward, and will provide a clearer picture for educators of the next generation of evaluation and development systems. As advocates for great public schools for every child, ConnCAN and 50CAN also hope that advocates can use this information as proof that we can—and must—move forward with smart, balanced, and fair evaluations of a teacher’s impact on student performance and growth. Such evaluation systems are fundamental to ensuring great teachers for every child. Because great schools, and great teachers, change everything.



INTRODUCTORY BRIEF

In the past, we have had, at best, a vague sense of how our teachers were performing and their impact on student learning. After just one or two observations, often lasting fewer than 80 minutes in all, most teachers received some sort of satisfactory rating—as did 94+ percent of their peers.¹ Afterward, we had no better idea of teacher quality, how to best develop teachers’ talents, or how to address ineffectiveness to support students’ academic needs.

Today, all of that is changing. States, school districts, charter school networks, and schools of education across the country are rethinking the way they measure teacher effectiveness, and using that information to improve teacher practice and student outcomes. As part of that effort, these systems are adopting new educator evaluations that include multiple measures, including teacher observation, student performance, student perceptions, community involvement, and other factors that foster the conditions students need to achieve at high levels. Sites are using these measures not only to differentiate between multiple levels of teacher effectiveness, but also as a tool to inform important decisions about teacher professional development, retention, dismissal, and pay, all of which are aimed at increasing overall teacher effectiveness and improving student outcomes. The best way to make these changes and others, however, is neither easy nor obvious. With this challenge in mind, ConnCAN and 50CAN approached Public Impact for this report.

REPORT GOALS

This report highlights 10 of the most advanced and talked-about teacher evaluation systems nationally: Delaware; Rhode Island; Tennessee; Hillsborough County, FL; Houston, TX; New Haven, CT; Pittsburgh, PA; Washington, DC (referred to throughout just as Washington); Achievement First (a charter management organization, or CMO); and the Relay Graduate School of Education in New York City. Together, these systems serve more than 1.6 million students each year. These are not the only systems taking on this work, but they are trailblazers. None of the sites we feature would say they have fully “figured out” teacher evaluation. Nor do we hold them up as examples of perfection. Yet they have all worked long and hard to carefully and thoughtfully tackle the most difficult challenges related to developing and implementing a high-quality teacher evaluation system.

As more states, districts, school systems, and schools of education strive to develop and implement teacher evaluation systems, they will wrestle with these same challenges. This report is designed to help share information and lessons learned about this difficult, yet essential, work. Our report therefore strives to “get under the hood” of the teacher evaluation systems at these 10 sites. We have collected information about these emerging systems in a single place, using a consistent format that allows users to look across sites and see where organizations are proceeding in similar ways, and where approaches diverge. Although we recognize that there are many uses for data related to teacher effectiveness, and that many sites are grappling with how best to use this information to improve teacher quality, this report focuses specifically on evaluation practices.

Other organizations, including the National Council for Teacher Quality, The New Teacher Project, the Aspen Institute, and Bellwether Education Partners have provided invaluable information about the emerging state of teacher evaluation policy; but these policies inevitably leave holes to be worked out during implementation. We do not know of any other report that collects as much

¹Weisberg, D, Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect*. The New Teacher Project. Available: <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>

detailed technical information as this report to describe how states, districts, and charter organizations are implementing these evaluation policies. We hope that this report serves as a “go-to” tool for policymakers working to reform an evaluation system or create one anew.

We must also note what this report does not do. It does not assess the merits of the evaluation systems; instead, it aims to describe the processes sites are using at a particular time and explain, when possible, why they decided to go in that direction. Nor does this report aim to include all of the possible options for addressing the challenges highlighted, as many other works have already contributed much to that conversation.² Instead, this report focuses on the methods used across the featured sites. The report also does not address how other school staff members are evaluated. Although we recognize that other school staff, especially the school leadership team, are central to reform, and that many sites have developed or are developing robust systems to evaluate them, a survey of that work went beyond the scope of this report.

REPORT COMPONENTS

This report consists of four components:

1. This **brief**, which provides an overview of the report and background on the topics addressed in the other documents, including key questions, implementation options and trade-offs, and key terms
2. A **cross-site analysis** that looks across all 10 sites and summarizes key components of their evaluation systems
3. **10 detailed profiles** of the teacher evaluation systems at our featured sites
4. A **library of documents** that are the building blocks of the 10 systems

These components build on one another, with each providing increasing detail. This brief offers a logical starting point for understanding key issues and terms and getting an overview of teacher evaluation in the 10 sites. The cross-site analysis provides a high-level overview of the sites side by side, enabling users to compare systems and key topics. The site profiles offer details on each of the sites, covering all of the aspects highlighted in the cross-site analysis and more. Finally, the library of documents offers links to documents and reports by and about the sites.

Because all of these sites continue to develop their evaluation systems, this work represents a snapshot in time and is current as of March 15, 2012. We plan to update this report over time to reflect new developments.

² See, for example: The New Teacher Project. (2010). *Teacher evaluation 2.0*. Available: <http://tntp.org/ideas-and-innovations/view/teacher-evaluation-2.0>; Public Impact. (2009.) *How should states define teacher effectiveness?* Available: http://www.publicimpact.com/publications/PublicImpact-How_Should_States_Define_Teacher_Effectiveness.pdf; National Comprehensive Center for Teacher Quality. *Database on teacher evaluation policies*. Available: <http://resource.tqsource.org/stateevaldb/>; Curtis, R., & Wiener, R. (2012, March). *Means to an end: A guide to developing teacher evaluation systems that support growth and development*. Washington, DC: The Aspen Institute Education and Society Program. Available: <http://www.aspeninstitute.org/publications/means-end-guide-developing-teacher-evaluation-systems-support-growth-development>

METHODOLOGY

We spent several months identifying sites, determining which information to collect, scouring public documents, conducting site interviews, soliciting feedback, and revising our materials. Below we describe each of the steps we took to create the documents included in this report:

- Select sites.** We set out to identify 10 sites to feature for this report. We sought sites that, while still evolving, were advanced enough to describe in detail. Our initial list consisted of about two dozen sites, including states that had recently adopted new legislation related to teacher evaluation or won Race to the Top grants; districts that had received grants through the Bill & Melinda Gates Foundation or the Teacher Incentive Fund to develop their teacher evaluation systems; sites that had recently been featured in other research, news reports, or education blogs; and sites recommended by experts with whom we consulted. Using that list, we conducted preliminary research to determine whether the evaluation system was in effect as of the 2011–12 school year, and to gauge how much information we could collect on the sites. Based on our research, we narrowed our list to 10 sites: three states, five districts, a charter school network, and a graduate school of education.

FEATURED SITES		
Delaware	Hillsborough County, FL	Achievement First (CMO)
Tennessee	Houston, TX	Relay Graduate School of Education
Rhode Island	New Haven, CT	
	Pittsburgh, PA	
	Washington, D.C.	

- Develop a framework for collecting the data.** We drew on recent research to identify the most common and vexing challenges sites face as they attempt to establish or revise their teacher evaluation systems. Next, we developed a series of questions related to those challenges. We also consulted with several experts to review and revise our questions, and we continued to tweak the framework as we gathered additional information or when more appropriate questions came to light.
- Review public documents.** For each site, we scoured publicly available documents describing the evaluation system and how it works. These included official documents from the site, as well as previous reports and news stories describing the evaluation system. Many of these documents are hyperlinked in the site profiles or available in the document library for this report.
- Conduct site interviews.** We interviewed at least one representative from each of the sites, and often interviewed two or more to answer remaining questions. These interviews included a mix of state and district personnel and consulting groups involved in designing and implementing the evaluation system.
- Review process.** After we completed the report documents and reviewed them internally, at least one representative from each site reviewed all of that site’s materials.

IMPLEMENTATION CHALLENGES

The rest of this brief describes major implementation challenges identified in our site research and interviews, and the methods sites are using to address those challenges. The challenges generally fell into five categories:

- Student achievement measures
- Classroom observations
- Other nonacademic measures
- Accuracy, validity, and reliability
- Reporting and using evaluation results

Our goal here is not to describe each category or challenge in full—which would take many more pages than we have here—but to highlight and discuss a set of critical issues that any district, state, charter organization, or school of education is likely to face. Throughout this brief, we therefore provide references to other works that dive into these issues in more detail.

STUDENT ACHIEVEMENT MEASURES

A teacher's primary job is to improve student learning. In the past, most school systems have not included student achievement as a factor in a teacher's evaluation. That tendency is changing, but fairly and accurately tying student growth to a particular teacher has proved difficult.³ To name just a few challenges: Student assessment data does not tell us everything about a teacher; many teachers teach untested grades or subjects; and teachers work with students starting from different points and facing different challenges.⁴ All of the sites profiled in this report have concluded, however, that including student achievement measures is worth the effort, because student learning progress is the core result for which all of the schools and districts strive. In addition, researchers have reached positive conclusions about common measures of teachers' contributions to student learning in core subjects, such as the correlation with students' mastery of higher-order skills and the year-to-year consistency of scores.⁵

Not all of the sites we featured have rolled out their systems' student achievement component yet, but they have all at least drafted a plan to do so. All of the sites have therefore wrestled with such tough questions as: how to measure student performance in untested grades and subjects; methods for calculating student growth; whether and how to adjust scoring for untested grades and subjects; and how heavily to weigh student performance in a teacher's final rating. Each site's decisions and experiences with these issues can offer lessons to others taking on the same challenges.

³Weisberg, D, Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect*. The New Teacher Project. Available: <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>

⁴For more, see: Stumbo, C., & McWalters, P. (December 2010/January 2011). Measuring effectiveness: What will it take? *The Effective Educator*. Vol. 68, No. 4. Available: <http://www.ascd.org/publications/educational-leadership/dec10/vol68/num04/Measuring-Effectiveness@-What-Will-It-Take%2%A2.aspx>; Goe, L., Bell, C., & Little, O. (2008, June). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Available: <http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf>

⁵The Bill & Melinda Gates Foundation. (2012). *Gathering feedback for teaching*. Available: http://metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf; Glazerman, S. et al. (2010). *Evaluating teachers: The important role of value-added*. Washington, DC: Brookings Institution. Available: http://www.brookings.edu/reports/2010/11/17_evaluating_teachers.aspx

MEASURING STUDENT ACHIEVEMENT IN UNTESTED GRADES AND SUBJECTS

Most of the sites we reviewed use the results from state standardized tests to measure student achievement when those data are available. Most teachers do not teach a class for which there is a state standardized test, however.

Several sites are expanding their student assessment options by creating new standardized assessments in untested grades and subjects. Pittsburgh Public Schools, for example, is using locally developed assessments to complement statewide tests, as well as working to expand its assessment portfolio over time. In Hillsborough County Public Schools, the district has built out its assessment portfolio to include tests for every grade and subject the district offers.

Other sites use standardized test results from a class or subject that the teacher's work supports. In Tennessee, for example, state law requires that student growth data count for 35 percent of a teacher's overall rating. For untested grades and subjects, the state has provided guidance on appropriate school-level value-added scores that districts should use.

Most sites, however, are choosing alternate measures of student achievement for untested grades and subjects, including the following:

- **Existing assessments.** Many of the sites are vetting and approving assessments used elsewhere. In grades K–2, for example, our featured sites consistently approved using several nationally normed literacy assessments for teacher evaluation, including the [Developmental Reading Assessment \(DRA\) 2](#) and the [Fountas and Pinnell benchmark assessment system](#).
- **School or teacher-created assessments.** Most sites allow teachers to create their own assessments if an approved assessment is not available.
- **Portfolio.** In some grades and subjects that are particularly difficult to test, sites are assessing student performance using a portfolio, which includes student work samples and other materials that demonstrate growth. At Achievement First, for example, teachers of art, music, physical education, dance, and theater decide on their individual program goals and assessments with a network achievement director at the beginning of the year. At the end of the year, a content expert reviews teachers' portfolios of student assessments to determine their progress toward their goals and their impact on student achievement.

In addition, several sites are using *multiple measures* to assess student achievement, even when standardized test data are available. According to those interviewees, the rationale is that no single measure is perfect, but combining multiple measures diminishes the weaknesses of any particular measure. In Houston, the district plans to require every teacher to have at least two measures of student performance by 2012–13. Similarly, Delaware plans to measure student growth using three measures—a state assessment-based measure, an approved external/internal assessment measure, and a growth goal measure. Meanwhile, Pittsburgh served as the pilot site for the Measures of Effective Teaching (MET) project through the Bill & Melinda Gates Foundation, and is considering whether and how to include different measures that research has identified as highly correlated with student achievement results.

METHODS FOR CALCULATING GROWTH

After sites collect student performance data, they must still calculate how much learning growth students make with a given teacher. Value-added models and student growth percentiles consider a student's prior performance to determine whether they are making as much growth as they should be:

- **Value-added model.**⁶ Value-added models use a statistical formula to predict how much growth a student will make based on the amount of growth that similar students—including students performing at a similar level at the start of the year—make, or have made historically. Some models even aim to parse out individual teacher effects among teachers who team-teach.⁷ A teacher's effectiveness is therefore measured in terms of the growth her students make compared to the amount of growth the model predicts that those students will make.

Several sites, including Tennessee, Washington, Pittsburgh, and Achievement First use teacher-level value-added data to assess student growth.⁸ Two sites, Tennessee and Washington, also include school-level value-added data, which looks at student results across the entire school, rather than at the teacher level. Pittsburgh is considering including a school-level value-added measure in teachers' summative ratings as well.

- **Student growth percentile.**⁹ A student growth percentile shows student progress in comparison to his academic peers—students whose performance was similar on previous assessments. Each child receives a percentile rank, indicating the percentage of his academic peers that he out-grew. Rhode Island, the only site we looked at that uses student growth percentiles to measure growth, uses the median (middle) Student Growth Score (percentile rank) to summarize student growth across a teacher's class. This number captures the point at which half the students had higher levels of growth and half the students had lower levels of growth. If the median is greater than 50, the teacher has exceeded expected growth, and vice versa.

Sites also used other methods for calculating growth that do not include comparable data, such as progress towards growth goals and mastery of standards:

- **Growth from pre-test to post-test.** Several sites set a growth target, then measure student performance against that target using a pre-test and post-test. In New Haven, teachers work with administrators to develop growth goals for their students; for example, students will grow, on average, at least four units on the elementary reading assessment from the beginning of the year to the end. Teachers are then evaluated on the progress students make towards those growth goals.

⁶For more on value-added models, see: McCaffrey, D., Lockwood, J., Koretz, D., & Hamilton, L. (2003). *Evaluating value-added models for teacher accountability*. RAND Corporation. Available: http://www.rand.org/pubs/monographs/2004/RAND_MG158.pdf; Mathematica Policy Research. *Using value-added growth models to track teacher and school performance*. Princeton, NJ: Author. Available: http://www.mathematica-mpr.com/education/value_added.asp; Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service. Available: <http://www.ets.org/Media/Research/pdf/PICVAM.pdf>

⁷Of the sites we reviewed, the value-added models used by Hillsborough County and Pittsburgh Public Schools aim to parse out individual teacher effects. In Rhode Island, the state weights teacher results to reflect the time each teacher spends with a particular student.

⁸Teacher-level value-added results do not currently contribute to a teacher's summative rating in Pittsburgh, but will beginning in the 2013–14 school year.

⁹For more on student growth percentiles, see: Betebenner, D. (2008). *A primer on student growth percentiles*. Dover, NH: National Center for the Improvement of Educational Assessment. Available: <http://www.cde.state.co.us/cdedocs/Research/PDF/Aprimeronstudent-growthpercentiles.pdf>

- **Mastery of standards.** In some grades and subjects, especially high school elective classes such as economics or journalism, growth goals are not appropriate because students enter the class with no or limited background in the subject. Some sites therefore evaluate teachers based on students' mastery of the standards instead. At Relay Graduate School of Education, teachers must demonstrate that their students, on average, have made at least a year's worth of growth or have mastered at least 70 percent of the grade-level standards to earn a degree.

Overall, our interviewees expressed that although student growth measures and our methods for calculating growth are imperfect, they remain useful and are constantly improving. In the long term, interviewees told us, they hope to find and create more reliable and valid assessments for non-tested grades and subjects that are independently administered and graded, but they acknowledge that such measures are not available in the short term. They also noted that the formulas underlying these scores must be robust, and the data systems in which test information is collected and analyzed must have the capacity to ensure the data are accurate and secure.

WEIGHTING STUDENT PERFORMANCE

At the end of the day, the evaluation systems we researched produce, or will soon produce, a summative score estimating the effectiveness of each teacher, which requires sites to combine all of the components by weighting them relative to their importance. In all of the systems we examined, student performance is a major factor in a teacher's final score. The sites differ, however, in how they weight student performance. Here are some of the approaches we saw:

- **Flat percentage.** Five sites use a flat weight for student performance. The weights fall between 40 percent (Hillsborough County and Achievement First) and 55 percent (Washington). Some sites lower that weight for teachers of untested grades and subjects, however. If value-added data are not available, student performance counts for just 15 percent of a teacher's final rating in Washington, and 20 percent at Achievement First.
- **Matrices.** Several sites are using or are planning to use a matrix that combines ratings on different scales into a single rating. New Haven, for example, rates each teacher on two factors: (1) student learning growth, and (2) instructional practice and professional values. A matrix then assigns a final rating to the teacher based on the combination of (1) and (2). For example, a teacher rated "strong" in growth and "effective" in practice and values would receive a final rating of "strong."
- **Rating ceiling.** Delaware places a ceiling on the final rating a teacher can receive based on her student's growth. Teachers must demonstrate satisfactory growth or better to be deemed effective or highly effective.
- **Providing a substitute measure.** At some sites, the evaluation system substitutes a teacher-level value-added score with another growth measure as needed. For example, if value-added data are not available for a particular teacher in Washington, that teacher receives a score for "teacher-assessed achievement data," a measure of student progress toward a target. Similarly, Houston will use teacher-level value-added data for any teacher for whom they are available, but has identified alternate measures for teachers of untested grades and subjects.

CLASSROOM OBSERVATIONS

In addition to measuring teacher effectiveness by using student performance data, all of the sites include classroom observation. There was a surprising amount of consensus around what to look for during observations. Most of the sites use Charlotte Danielson's *Framework for Teaching* to assess teacher practice, or have built their own rubric based on the Danielson model, which looks across four domains of teaching: planning and preparation, the classroom environment, instruction, and professional responsibilities.¹⁰ Who conducts those observations and how often teachers are observed varies from site to site.

OBSERVERS

The sites we spoke to identified several considerations related to choosing observers.

- Who understands the school and classroom context in which the teacher works?
- Who has the subject matter and grade-level expertise to provide a valid and meaningful evaluation?
- Who has the time to conduct all of the evaluations needed?
- Who is sufficiently objective to follow a rubric reliably?
- What approaches are affordable within our budgets?

Based on these questions, the sites generally chose among administrators, a third-party evaluator unaffiliated with school, and teacher peers within the school to serve as observers. As we describe below, each type of evaluator has pros and cons, so a particular type may be more appropriate in a specific situation:

- **Administrators and coaches.** Administrators and coaches see teachers every day and understand the conditions under which teachers work. Administrators may not have subject matter expertise, however, and may not have enough time in their schedules to get into the classroom and evaluate teachers as needed. And they may have trouble making objective determinations, especially as they get to know teachers over time. Administrators serve as observers in every school system we reviewed.
- **Third-party evaluators unaffiliated with the schools.** Third-party evaluators are usually subject-matter and grade-level experts with a history of high performance in the classroom. Their primary job is to go from school to school conducting teacher evaluations. They have the benefit of being impartial, because they do not know the teacher or the school, and their results serve as a check on results from other observers. But third-party evaluators also lack an understanding of the school and the special challenges a particular classroom may face. Third-party evaluators are paid for this work, increasing total evaluation costs. Third-party evaluators are a key component of several systems we reviewed, including Hillsborough County, Washington, and New Haven.

¹⁰ See the cross-site analysis and individual fact sheets for examples where sites did not use Charlotte Danielson's *Framework for Teaching*.

- **Peers within the school.** Like administrators, other teachers in the school understand the conditions under which the teacher works. They are also likely to have subject-matter expertise. While some districts have found peers to be “tougher” evaluators than administrators, peers arguably may have trouble making objective determinations when they have close relationships with the teachers being evaluated.¹¹ Both Pittsburgh Public Schools and Achievement First use peer evaluators from within the school, but their observations do not factor into teachers’ summative ratings.¹²

Sites do not necessarily use just one type of observer. Although administrators conducted observations at every site, sites sometimes used Third-party evaluators as well, allowing teachers to be observed more frequently when needed. At Achievement First, for example, where internal school leaders observe alongside a network instructional expert (regional superintendents or an achievement director with expertise in a particular content area), the use of multiple observers also ensures adherence to a network-wide standard of excellence and provides meaningful professional development for school leaders.

FREQUENCY OF OBSERVATIONS

Some sites observed all teachers an equal number of times; other sites observed teachers with varying frequency based on what would be most needed or helpful, given limited resources and a teacher’s evaluation results. In some systems, this broke down by the following types of teachers:

- **Low-performing teachers.** Additional observations can be a way to provide additional support for low-performing teachers. Such is the case in Delaware, where the average teacher receives just one observation annually, but low-performing teachers receive two.
- **High-performing teachers.** The frequency with which high-performing teachers were observed varied across sites. In New Haven, both low-performing and high-performing teachers receive additional observations. A third party evaluator conducts an additional observation to verify that the teacher truly performs at a high level, because such teachers are eligible to become a lead teacher, mentor, or coach, and may be called upon to help other teachers improve. In other systems, such as Washington’s, high-performing teachers have the option to receive fewer observations as a way to recognize some of the district’s best teachers and allow principals to spend more time supporting struggling teachers.
- **Novice teachers.** Additional observations can help provide more opportunities for feedback and development early in a teacher’s career, when data indicate teachers achieve the most professional growth.¹³ In Tennessee, for example, novice teachers receive six observations per year, while other teachers receive four. Interestingly, none of the sites we reviewed require fewer observations for tenured teachers.

¹¹ Papay, J., & Johnson, S. (2011). Is PAR a good investment? Understanding the costs and benefits of teacher peer assistance and review programs. *NGT Working Paper*. Available http://www.gse.harvard.edu/~ngt/new_papers/PAR%20Costs%20and%20Benefits%20-%20January%202011.pdf

¹² Teachers filling the Instructional Teacher Leader 2 (ITL2) position at PPS will conduct observations at their own schools in their first year as an ITL2, and those observations will be used for formative purposes, only. In their second and third years, ITL2s will observe teachers in other schools, serving more as third party evaluators, and those observations will be factored into teachers’ summative ratings. See fact sheet for more detail. At Achievement First schools, peer observations do not factor into a teacher’s final observation score.

OTHER NONACADEMIC MEASURES

Many of the sites we researched use other nonacademic measures in addition to classroom observation. By using a combination of academic and nonacademic measures, they aim to capture different aspects of effective teaching. Some of the other nonacademic measures sites are using or exploring include:

- **Student perceptions.** This year, Pittsburgh began a district-wide administration of the Tripod student survey, which measures a variety of nonacademic student outcomes. Preliminary results show a strong correlation between some elements of students' perceptions demonstrated on these surveys and teachers' value-added scores. The district has not yet decided, however, how and whether to use the results as a factor in teachers' overall ratings. Washington is also running a small pilot of the Tripod survey, but has no immediate plans to include student surveys as part of its teacher evaluation system. Rather, the district hopes to make the results available to teachers as a tool for improvement.
- **Student character.** Both Achievement First and Relay Graduate School of Education score teachers based on growth in student character. At Achievement First, the student character component is based on student and parent surveys about relationships and communication with students and families.
- **Peer ratings.** At Achievement First, a teacher receives a score for "core values and contribution input," which is based on a peer survey that assesses a teacher's core values and contribution to the AF mission.
- **Commitment/contribution to school community.** Every site we looked at included a measure of professionalism, such as collegiality and high expectations for students. Washington, however, includes an extra component: 10 to 15 percent of a teacher's final rating reflects her commitment and contribution to the school community. Administrators measure that commitment using a rubric that scores a teacher's support of local school initiatives, support of special education and English language learner programs, high expectations, partnership with families, and instructional collaboration.

VALIDITY AND ACCURACY OF DATA

All of the sites recognized that teacher evaluation results are valuable only if the measures used are accurate and valid. Sites are generally early in their efforts to assess data accuracy and validity, but are taking some steps to do so, including:

- **Evaluator training.** All of the sites train their evaluators. The length and intensity of the training differ by site, however, as do the amount and types of ongoing support. In Houston, for example, evaluators participated in a four-day training session over the summer this year. In Hillsborough County, training lasted for seven to 10 days and included conducting

¹³ See for example: Nye, B., Konstantopoulou, S., Hedges, L. (2004). "How Large Are Teacher Effects." *Educational Evaluation and Policy Analysis*. Vol. 26, No. 3. Pp 237-257. Available <http://steinhardt.nyu.edu/scmsAdmin/uploads/002/834/127%20-%20Nye%20B%20%20Hedges%20L%20%20V%20%20Konstantopoulou%20S%20%20%20%282004%29.pdf>; Clotfelter, C., Ladd, H., Vigdor, J. (2007). "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." Calder Center. Available Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects," Calder Center, October 2007; Harris, D., Sass, T (2007), "Teacher Training, Teacher Quality, and Student Achievement," Calder Center. Available http://www.caldercenter.org/PDF/1001059_Teacher_Training.pdf

paired observations with another evaluator. In Pittsburgh, all evaluators participate in the Instructional Quality Assurance Certification (IQA-C) Process, which includes two certification levels focusing on rating accuracy and instructional feedback and support.

- **Evaluator tests.** At several sites, including Delaware, Tennessee, and Houston, evaluators must pass a test before they can evaluate teachers.
- **Evaluators' ratings.** Evaluators are themselves evaluated, based in part on the accuracy and validity of their ratings. This is most common at sites that use third-party evaluators.
- **Third party evaluators used.** The use of a third-party evaluator serves as a check on administrator ratings in the building.
- **Procedure to investigate misalignment between measures.** In the case that the score on different measures, such as student performance and observation, are significantly different, several sites will conduct a review. In Rhode Island, for example, the district will review the evaluation data for a particular teacher if a teacher has an extremely high score for student achievement, but an extremely low score for professional practice and responsibilities, adjusting as needed.
- **Regular data checks and follow-ups.** In Rhode Island and a handful of other sites, district officials are continually monitoring evaluation data. In instances when the data seem “off,” such as results from a particular school or evaluator being very different from others, district officials will share the results with those involved to determine the cause of the discrepancy.
- **Reviews of student performance results.** Every year, Relay Graduate School of Education randomly reviews the student performance results of about 5 to 10 percent of its teachers. During the review, Relay GSE reviews all documentation and has a conference with the teacher about his or her students' work.
- **Principal discretion combined with superintendent sign-off.** At Achievement First, teachers wanted principals to have more discretion, rather than less. Principals can review student data and use other evidence of teacher performance to add points to the “student achievement outcome” component. In those instances, a regional superintendent reviews the principals' decisions.

REPORTING AND USING EVALUATION RESULTS

REPORTING RESULTS

Once teacher evaluations are complete, sites must decide what to do with the results. At every site, teachers and school leaders received their scores. Most sites also made public the distribution of ratings across all schools in the system. For example, Washington released the percentage of teachers falling into each performance category. In Rhode Island, the state department of education also plans to publish report cards for all educator preparation programs in the state based on teacher evaluation results beginning at the end of the 2014–15 school year.¹⁴

There seems to be little effort or plan to report scores in any further detail. Sites continue to struggle with the question of how to share evaluation information, but generally hesitate to report the data in a way that makes it possible for the public to link results to a specific teacher.

USING RESULTS

All the sites in this report use teacher evaluation results to incentivize and reward different behaviors and outcomes that pay off for students, and to help administrators make decisions about individual teachers to improve a school's performance. Across the sites, states, districts, and schools are using the results to determine:

- **Professional development.** Almost all of the sites reported using teacher evaluation data to create targeted professional development. Several sites also reported using evaluation results to identify teachers for an improvement plan.
- **Bonuses and salary increases for top performers.** Based on their performance, effective teachers at Achievement First schools can earn significantly more than current salary scales within the surrounding traditional public school districts. In Washington, teachers deemed “highly effective” can receive bonuses of up to \$25,000. If they earn “highly effective” ratings for multiple years, they can increase their base salary by as much as \$27,000.
- **Promotion/increased responsibility.** Several systems offer or plan to offer opportunities for teachers to take on additional responsibility or new roles. In Hillsborough County, for example, top-performing teachers can apply to become a peer evaluator. In New Haven, exemplary teachers are eligible to be a lead teacher, mentor, or coach.
- **Student assignment.** Rhode Island plans to implement a policy in which no student will have a low-performing teacher two years in a row.
- **Tenure.** In Tennessee, new teachers can receive tenure only if they teach in the district for five years and attain a rating in the top two evaluation categories for the previous two years. In Pittsburgh, teachers must earn six satisfactory ratings before receiving tenure.
- **Licensure/Certification.** In Delaware, new teachers must earn a satisfactory rating at least twice within three years to earn a continuing license. Similarly, teachers enrolled at Relay Graduate School of Education cannot earn a degree unless they demonstrate measureable growth in the classroom. And in Rhode Island, even experienced teachers will not be able to renew their certification if they are deemed ineffective for five years. Rhode Island is also the only site that uses differentiated certification (initial, professional, and advanced) based on evaluation results.

¹⁴ Louisiana has long been the leader in matching teacher evaluation results to teacher preparation programs. See, for example: <http://www.regents.doa.louisiana.gov/assets/docs/TeacherPreparation/RegentsReceipt11FINAL.pdf>

- **Dismissal.** Several sites consider evaluation results in dismissal decisions. In Hillsborough County, for example, teachers who receive a rating of 1 or 2 for two consecutive years could be recommended for dismissal. Teacher dismissal procedures in Hillsborough County and several other sites we reviewed, however, are often set by state statute and have not changed as a result of new evaluation systems.

CONCLUSION

The additional sections of this report provide more detail on the 10 evaluation systems we reviewed. As we mentioned at the start of this brief, none of these systems claims to have cracked the code for teacher evaluation. Nonetheless, we consistently heard that the perfect should not be the enemy of the good. As one interviewee described his system's latest student performance measure: "This is an imperfect measure. It's the best one we ever had."

Although they are still works in progress, the evaluation frameworks in place at these 10 sites successfully use multiple measures to gauge teacher effectiveness. Gathering data on teacher effectiveness is only the first step, however. What sites do with the data is just as critical. The sites we reviewed use the data they collect to differentiate teacher quality, which allows them to reward excellence, remove poor performers, use talent more effectively, provide targeted professional development, and elevate the teaching profession. Most important, these evaluation systems and the measures collected provide a key instrument for improving student achievement for more than 1.6 million children every year.



CROSS-SITE ANALYSIS

This section of *Measuring Teacher Effectiveness* looks across all 10 sites and summarizes key components of each evaluation system. It provides a high-level overview of the sites to enable side-by-side comparisons and allow readers to zero in on the most relevant sites and topics. For more detail on any of the systems, please see the site profiles.

The tables in this section compare and contrast the following major components of the teacher evaluation systems we reviewed:

- Background on each evaluation system
- Student achievement measures
- Classroom observations
- Other nonacademic measures
- Accuracy, validity, and reliability
- Reporting and using evaluation results

		STATES			DISTRICTS					OTHER		
		DELAWARE	RHODE ISLAND	TENNESSEE	HILLSBOROUGH COUNTY, FL	HOUSTON, TX	NEW HAVEN, CT	PITTSBURGH, PA	WASHINGTON, D.C.	ACHIEVEMENT FIRST (CMO)	RELAY GRADUATE SCHOOL OF EDUCATION	
BACKGROUND ON EVALUATION SYSTEMS	Number of students	130,610	143,793	933,703	192,547	202,773	20,759	25,000	45,000	6,200	21,000	
	Number of schools	208	300	1,736	250	298	45	60	125	20	93	
	Number of teachers	8,594	14,260	64,229	12,468	12,829	1,850	2,000	4,000	585	420	
	Most recent evaluation system implemented systemwide in...	2011-12	2011-12	2011-12	2010-11	2011-12	2010-11	2010-11	2009-10	2010-11	2011-12	
	Other staff evaluated using an aligned process	Principals	X	X	X	X	X ¹⁵	X	X ¹⁶	X	X	n/a
		Non-teaching staff	X		X	X		X		X		n/a
	How often are teachers evaluated?	Teachers in general	1x/2yrs	1x/yr	1x/yr	1x/yr	1x/yr	1x/yr	1x/yr	1x/yr	1x/yr	Once at end of year 2
		High-performing teachers	1x/2yrs	Evaluation intervals do not change based on teacher performance								n/a
		Low-performing teachers	1x/yr	Evaluation intervals do not change based on teacher performance								n/a
		Novice teachers	1x/yr	Evaluation intervals do not change based on tenure / experience					2x/yr	Evaluation intervals do not change based on tenure/ experience		n/a
Tenured/ experienced teachers		Varies (1x/1-2 yrs)	Evaluation intervals do not change based on tenure / experience					1x/yr	Evaluation intervals do not change based on tenure/ experience		n/a	

¹⁵Houston is currently designing an aligned principal appraisal system.

¹⁶Pittsburgh's principal evaluation system is being updated to align with the district's teacher evaluation system.

			STATES			DISTRICTS					OTHER			
			DELAWARE	RHODE ISLAND	TENNESSEE	HILLSBOROUGH COUNTY, FL	HOUSTON, TX	NEW HAVEN, CT	PITTSBURGH, PA	WASHINGTON, D.C.	ACHIEVEMENT FIRST (CMO)	RELAY GRADUATE SCHOOL OF EDUCATION		
STUDENT ACHIEVEMENT MEASURES	Student assessments used for evaluation	Standardized state assessment	X	X	X	X	X	X	X	X	X			
		Existing assessment	X	X	X	X	X	X	X	X	X	X		
		School- or teacher-created	X	X	X		X	X		X	X	X		
		Portfolio	X		X		X	X		X	X			
	What is the student achievement measure for tested grades/subjects?	Teacher-level value-added measure				X	X			X	X	X		
		Student growth percentile		X										
		Growth from pre-test to post-test	X	X			X	X					X	
		Mastery of standards	X ¹⁷	X	X								X	
		Other (e.g., achievement)												
	What is the student achievement measure for untested grades/subjects? ¹⁷	Value-added measure from a related subject or grade				X	n/a all students are tested			TBD				
		Student growth percentile		X										
		Growth from pre-test to post-test	X	X	X ¹⁸			X	X			X	X	X
		Mastery of standards	X ¹⁹	X	X			X				X		X
		Other (e.g., teacher-assessed achievement data)				X						X		
Does the system include a school-level student achievement measure?				X				TBD	X					
Weight of student achievement component in tested grades/subjects		Varies / Uses Rating ceiling	Varies / Uses matrix	50%	40%	Varies / Uses matrix	Varies / Uses matrix	TBD	55%	40%	45%			
Weight of student achievement component in untested grades/subjects		Varies / Uses Rating ceiling	Varies / Uses matrix	50%	n/a	Varies / Uses matrix	Varies / Uses matrix	TBD	15%	20%	45%			

¹⁷ Represents all possibilities, although not all measures are used for all teachers. See individual site profiles for more detail.

¹⁸ Currently being piloted for untested grades and subjects.

¹⁹ The Delaware Department of Education plans to look at external assessments that include proficiency/attainment until enough data is gathered to calculate growth.

			STATES			DISTRICTS					OTHER		
			DELAWARE	RHODE ISLAND	TENNESSEE	HILLSBOROUGH COUNTY, FL	HOUSTON, TX	NEW HAVEN, CT	PITTSBURGH, PA	WASHINGTON, D.C.	ACHIEVEMENT FIRST (CMO)	RELAY GRADUATE SCHOOL OF EDUCATION	
CLASSROOM OBSERVATIONS	Who are the observers who rate teacher effectiveness?	Administrators or coaches	X	X ²⁰	X	X	X	X	X	X	X		
		Third-party evaluators unaffiliated with school				X		X	X	X	X		
		Peers within the school											
		Other (e.g., regional superintendent, achievement director)			X						X	X	
	How do teachers receive feedback after an observation?	Conference	X	X	X	X	X	X	X	X	X	X	X
		Written feedback	X	X	X	X	X	X	X	X	X	X	X
	How often are teachers observed ²¹	Teachers in general	1x/yr	4x/yr	4x/yr	5x/yr	4x/yr	1x/yr ²²	4x/yr	5x/yr	3x/yr ²²	2-3x/yr	
		Low-performing teachers	2x/yr		4x/yr	>5x/yr ²³		4x/yr ²²	4x/yr	5x/yr			
		High-performing teachers	1x/yr		4x/yr	5x/yr		3x/yr ²²	4x/yr	2x/yr			
		Novice teachers	3x/yr		6x/yr	n/a		1x/yr ²²	8x/yr	5x/yr			
		Tenured teachers	Varies		4x/yr	n/a		Varies	4x/yr	5x/yr			
	Are formal observations announced?	Yes				X						X	
		No					X ²⁴				X		
		Sometimes	X	X	X			X	X	X			
What is the basis for the observation rubric?	Charlotte Danielson	X			X			X			X		
	Other (e.g. TAP, Teaching as Leadership, etc.)			X					X	X	X		

²⁰ Although most formal observations are conducted by principals or assistant principals, LEAs may determine who primary evaluators will be, as long as they receive training.

²¹ These groupings are the authors' creation. Sites may differentiate observations based on different teacher categories or terms.

²² Includes formal observations only. Number of informal observations not set, but "frequent."

²³ Hillsborough County, FL, conducts up to 11 observations, including 7 formal and 4 informal.

²⁴ Houston refers to observations as "longer" and "shorter" rather than "formal" and "informal."

		STATES			DISTRICTS					OTHER	
		DELAWARE	RHODE ISLAND	TENNESSEE	HILLSBOROUGH COUNTY, FL	HOUSTON, TX	NEW HAVEN, CT	PITTSBURGH, PA	WASHINGTON, D.C.	ACHIEVEMENT FIRST (CMO)	RELAY GRADUATE SCHOOL OF EDUCATION
What additional measures are included?	Student perceptions			X ²⁵				X ²⁶		X	
	Student character									X	X
	Peer ratings									X	
	Other (e.g., completion of class modules, contribution to school community—see profiles for details)		X						X		X

²⁵ Piloted this year in Memphis, but will expand next year.

²⁶ Pittsburgh Public Schools is administering a district-wide student survey this year, but has not committed to using the results to determine a teacher's evaluation score.

			STATES			DISTRICTS					OTHER		
			DELAWARE	RHODE ISLAND	TENNESSEE	HILLSBOROUGH COUNTY, FL	HOUSTON, TX	NEW HAVEN, CT	PITTSBURGH, PA	WASHINGTON, D.C.	ACHIEVEMENT FIRST (CMO)	RELAY GRADUATE SCHOOL OF EDUCATION	
ACCURACY, VALIDITY, & RELIABILITY	What steps is the site taking to ensure that results are accurate and valid?	Evaluator training	X	X	X	X	X	X	X	X	X	X	
		Evaluator tests	X		X	X	X		X				
		Evaluator ratings			X	X					X	X	
		Third-party evaluators used				X		X	X	X	X	X	
		Procedure to investigate misalignment between measures		X					X	X		X	X
		Regular data checks and follow-ups (e.g. district/school follow-ups)		X	X	X	X			X	X	X	
		Review of student performance results		X	X								X
		Other (e.g., peer survey, principal discretion)				X				X		X	

		STATES			DISTRICTS					OTHER		
		DELAWARE	RHODE ISLAND	TENNESSEE ²⁷	HILLSBOROUGH COUNTY, FL	HOUSTON, TX	NEW HAVEN, CT	PITTSBURGH, PA	WASHINGTON, D.C.	ACHIEVEMENT FIRST (CMO)	RELAY GRADUATE SCHOOL OF EDUCATION	
REPORTING AND USING RESULTS	How many rating categories are there?	4	4	5	5	4	5	4	4	5	n/a (scale 1-100)	
	How are results reported and used?	Teachers and administrators receive score	X	X	X	X	X	X	X	X	X	X
		Results are used to evaluate teacher preparation programs			X		X					X
	Results influence...	Professional development	X	X	X	X	X	X	X	X	X	n/a ²⁹
		Bonuses and salary increases for top performers	X		X	X			X	X	X	n/a ²⁹
		Promotion/increased responsibility		X	X	X		X	X	X	X	n/a ²⁹
		Student assignment		X	X							n/a ²⁹
		Tenure	n/a ²⁸		X		n/a ²⁸		X	n/a ²⁸	n/a ²⁸	n/a ²⁹
		Licensure	X	X	X							X ²⁹
	Dismissal		X	X	X	X	X		X			
Other (e.g., commendation, school-based appreciations, special recognition, etc.)	X		X			X		X	X	X ²⁹		
Teachers can appeal their ratings	X	X	X	X	X	X	X	N/A ³⁰	³¹	X	X	

²⁷The First to the Top Act requires that annual evaluations be a factor in personnel decisions, including promotion, retention, tenure, and compensation. The details of such decisions, however, are left to the district except with respect to tenure.

²⁸Delaware does not award "tenure," so teachers are either considered "novice" or "experienced" (holding a valid and current continuing or advanced license). Similarly, Houston, Washington, D.C., and Achievement First do not award "tenure."

²⁹Relay GSE does not issue a teaching license, but recommends teachers for certification if they complete the Relay GSE program. In addition, teachers receive a degree only if they earn a high enough score. All other "n/a"s to not apply to Relay GSE because it is a graduate school of education.

³⁰The district and union have appeal processes.

³¹Through the chancellor's appeals process, District of Columbia Public Schools provides teachers who earn ineffective or minimally effective ratings with an opportunity to submit an appeal if they do not believe the evaluation procedures were followed appropriately (e.g. a post-observation conference was not provided within 15 days). Teachers may appeal the evaluation process only, rather than observation scores themselves. The Washington Teachers' Union also provides teachers with an opportunity to submit appeals.



SITE PROFILES

DELAWARE

DELAWARE PERFORMANCE APPRAISAL SYSTEM FOR TEACHERS (DPAS II-REVISED)

OVERVIEW

Delaware’s Performance Appraisal System (DPAS) has undergone many changes since first being piloted in 1986. DPAS I, in effect until 2006–07, evaluated teachers on four components: planning and preparation, classroom environment, instruction, and professional responsibilities. DPAS II, implemented in 2007–08, added student growth as a fifth component in teacher evaluations. Then in 2010, Delaware’s Race To The Top application outlined revisions strengthening the student growth component.

In the state’s current evaluation system (DPAS II-Revised), the first four components of the evaluation are based on teacher observation. Delaware uses expert evaluators from the state and development coaches—former administrators and teacher leaders—to train administrators, calibrate observation results, and model observations and conference skills. Planning and preparation, classroom environment, instruction, and professional responsibilities have equal weight in a teacher’s rating, but a teacher cannot be deemed effective or highly effective if she does not produce at least a satisfactory rating on student growth, regardless of her ratings for the other components.

The Delaware Department of Education (DDOE) continues to work with stakeholder groups to refine the student growth component, including determining how much growth teachers must make to be deemed effective. In its current interim plan, DDOE uses multiple measures to evaluate student growth and provides teachers with options for calculating each measure. DDOE is also convening working groups of teachers to develop pre- and post-assessments for all grades and subjects.

ABOUT THE SCHOOL SYSTEM

<i>Number of students</i>	130,610
<i>Number of schools</i>	20,759
<i>Number of teachers</i>	8,594

BACKGROUND

<i>Corresponding legislation</i>	Race To The Top (RTTT) application revised DPAS-II
<i>Year legislation passed</i>	2010
<i>Was the system piloted?</i>	Yes
<i>In effect since</i>	<ul style="list-style-type: none"> • DPAS II: 2007–08 • DPAS II-Revised: Interim year 2011–12 (includes revised student growth component, but does not factor into overall rating; also ties rewards to evaluation results)
<i>Are there plans for additional phases / components?</i>	<ul style="list-style-type: none"> • 2011–12 is considered an “interim year,” and does not include the student growth component • Beginning in 2012–13, both positive and negative consequences will be tied to evaluation results, which will include the student growth component
<i>Who gets evaluated?</i>	All part-time and full-time teachers, specialists, and administrators

SITE PROFILES

<p><i>How often do teachers receive a rating?</i></p>	<ul style="list-style-type: none"> • Novice teacher: Summative evaluation once a year • Experienced teachers with most recent overall rating of highly effective: minimum of one summative evaluation every two years • Experienced teachers with most recent overall rating of effective, plus satisfactory ratings on at least four of the five evaluation components: minimum of one summative evaluation every two years • Experienced teachers who have not earned overall ratings of highly effective or effective: once-a-year summative evaluation <p>Note: Delaware does not award “tenure,” so teachers are either considered “novice” or “experienced” (holding a valid and current Continuing or Advanced License, respectively).</p>
<p><i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i></p>	<ul style="list-style-type: none"> • Yes, all educators must be evaluated annually using the DPAS II-Revised evaluation tool. However, LEAs may administer other evaluations in addition to DPAS II-Revised, as long as these are reported to DDOE • If an LEA wishes to use a locally developed evaluation system that is the result of a collective bargaining process, it may apply for a waiver of the DPAS II-Revised provisions. The DDOE will review the system to ensure that it: <ul style="list-style-type: none"> - is educationally sound and rigorous - measures student growth using multiple measures - has mechanisms in place to certify evaluators and validate results - does not apply to educators holding an initial license • The state also mandates that all districts use a state-created assessment system (DCAS), but enables districts to propose other models that meet state requirements such as: <ul style="list-style-type: none"> - LEA-wide diagnostic assessments (such as Reading inventory, Gates McGinitie) - LEA-wide common assessments (such as written curriculum-based assessments, performance-based assessments, student portfolios)
<h3>STUDENT ACHIEVEMENT MEASURES</h3>	
<p><i>Assessments to measure student achievement</i></p>	<ul style="list-style-type: none"> • Delaware Department of Education’s (DDOE) State assessment: Delaware Comprehensive Assessment System (DCAS) <ul style="list-style-type: none"> - DCAS math and reading assessments for grades 3–10 <ul style="list-style-type: none"> » Frequency: three times a year (fall, winter, spring) - DCAS Alternate Assessment (DCAS-Alt1)
<p><i>Are any new assessments in development to use with teacher evaluations?</i></p>	<p>Delaware is developing the internal measures that can be used by districts for the student growth component of the evaluation system. These measures must be approved by the Department of Education, and include:</p> <ul style="list-style-type: none"> • Internal pre- and post-assessments for non-DCAS content/subject teachers who report student grades • Growth goals for non-DCAS content/subject teachers who do not report student grades

<p><i>What academic/achievement criteria are included?</i></p>	<p>The student growth component is still in development. In 2011–12, only DCAS-based student growth will be calculated for DCAS-tested educators to determine highly effective educators. The 2012–13 school year plan for measuring student growth includes a DCAS measure, internal or external measures, and growth goals.</p> <ul style="list-style-type: none"> • DCAS Measure <ul style="list-style-type: none"> - Student DCAS fall-spring growth score (called “instructional score”) in reading and/or math for all of teacher’s students - This measure applies to all DCAS subject and grade teachers (grades 3–10, reading and math) • Internal or External Measures <ul style="list-style-type: none"> - Includes two options: <ul style="list-style-type: none"> » External measures: Non-DCAS assessments <u>approved by the DDOE</u> that are standards-based and show fall-to-spring growth³² <ul style="list-style-type: none"> › Examples: district-wide Measures of Academic Progress (MAP assessments) or IEP progress › Internal measures: statewide pre- and post-assessment created by educators <ul style="list-style-type: none"> › All are approved by the DDOE and meet criteria of <u>internal measures rubric</u> - Measures are selected by teacher and evaluator between August and October of the school year - This measure applies to all DCAS subject and grade teachers and non-DCAS teachers who give grades (e.g., science or social studies teachers) • Growth Goals <ul style="list-style-type: none"> - Each educator uses 15 indicators to measure growth, which include statewide and district-wide indicators that measure what educators are doing to assist in academic growth of students based on available data, job descriptions, and standards - This measure applies to non DCAS-subject teachers and teachers who do not give student grades (e.g., counselors and nurses)
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<ul style="list-style-type: none"> • Beginning in 2012–13, teachers will be able to earn an “effective” rating or better only if their students make satisfactory growth or better (an amount to be determined) • The various growth measures of student achievement will be weighted in the following ways: <ul style="list-style-type: none"> - For DCAS subject and grade teachers: <ul style="list-style-type: none"> » DCAS Measure: 50% » Internal or External Measures: 50% - For non-DCAS subject teachers: <ul style="list-style-type: none"> » Internal or External Measures: 50% » Growth Goals: 50% - For non-subject teachers who do not give student grades: <ul style="list-style-type: none"> » Growth Goals: 100%

³²DDOE plans to look at external assessments that include proficiency/attainment until enough data is gathered to calculate growth.

SITE PROFILES

NONACADEMIC MEASURES	
<p><i>What nonacademic evaluation criteria are included?</i></p>	<p>There are four nonacademic components in DPAS II-Revised: Planning and preparation, classroom environment, instruction, and professional responsibilities (see p. 7-34 of guide).</p> <ul style="list-style-type: none"> • Component 1: Planning and Preparation <ul style="list-style-type: none"> - Criteria are: selecting instructional goals; designing coherent instruction; demonstrating knowledge of content and pedagogy; demonstrating knowledge of students; and designing student assessments • Component 2: Classroom Environment <ul style="list-style-type: none"> - Criteria are: managing classroom procedures; managing student behavior; creating an environment to support learning; and organizing physical space • Component 3: Instruction <ul style="list-style-type: none"> - Criteria are: engaging students in learning; demonstrating flexibility and responsiveness; communicating clearly and accurately; using questioning and discussion techniques; and using assessment in instruction • Component 4: Professional Responsibilities <ul style="list-style-type: none"> - Criteria are: communicating with family; recording data in a student record system; growing and developing professionally; and reflecting on professional practice
<p><i>How much do the nonacademic measures count in a teacher's final rating?</i></p>	<p>Nonacademic components are weighted equally. Teacher receives one summative rating for these components (see p. 79 and 92 of guide) of satisfactory or unsatisfactory.</p>
CLASSROOM OBSERVATIONS	
<p><i>Who are the observers who rate teachers' effectiveness?</i></p>	<p>Certified evaluators: Usually principals or other administrators in the school administration.</p>
<p><i>What training do observers receive?</i></p>	<ul style="list-style-type: none"> • All evaluators must complete four training modules online, and pass all online quizzes, which total approximately six hours. If an evaluator does not pass a quiz after three tries, he will not be certified for the upcoming year • Evaluators earn a certification from DDOE that is valid for five years, and renewable upon completion of DDOE's DPAS II-Revised professional development requirements • Expert evaluators and development coaches receive further training from DDOE and the Delaware Academy for School Leadership (DASL). <p>For more information on administrator, evaluator, and teacher training, see here.</p>

<p><i>Are observers compensated for this work?</i></p>	<p>No, observations are part of the administrator's job description/salary.</p>
<p><i>How often are teachers observed?</i></p>	<ul style="list-style-type: none"> • New teachers: minimum of 3x/year • Experienced teachers with most recent overall rating of highly effective: minimum of 1x/year • Experienced teachers with most recent overall rating of effective, plus ratings of satisfactory on at least four of the five components: minimum of 1x/year • Experienced teachers who are not rated highly effective or effective: minimum of 2x/year
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<p>Observations occur after the first five full school days of the year, and before the last five days of the school year.</p> <p>Suggested target dates for completion of observations:</p> <ul style="list-style-type: none"> • New teachers: <ul style="list-style-type: none"> - Observation 1: by October 31 - Observations 2 and 3: by March 31 • Experienced teachers: <ul style="list-style-type: none"> - Observation: by January 31
<p><i>Are observations announced or unannounced?</i></p>	<ul style="list-style-type: none"> • New teachers: two announced and one unannounced • Experienced teachers rated highly effective or effective: one announced • Experienced teachers not rated highly effective or effective: one announced, one unannounced
<p><i>What is the basis for the observation rubric?</i></p>	<ul style="list-style-type: none"> • DDOE based the observation rubric on Charlotte Danielson's <i>Framework for Teaching</i>.
<p><i>What is included in the observation scoring rubric?</i></p>	<p>Evaluators use the <u>Components rubrics (pp. 114-131 of guide)</u>, based on <u>Charlotte Danielson's <i>Framework for Teaching</i></u>, during observations for nonacademic components:</p> <ul style="list-style-type: none"> • Component 1: Planning and Preparation • Component 2: Classroom Environment • Component 3: Instruction • Component 4: Professional Responsibilities
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • Administrators go through an online training program • DDOE is developing a credentialing process
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • All announced observations must be preceded by a pre-observation conference • All observations are followed by a post-observation conference within 10 days of the conference. Evaluators give <u>formative feedback (see p. 70 of guide)</u> with expectations and recommendations for teacher improvement • All teachers participate in a summative evaluation conference that discusses all previous observations

SITE PROFILES

ACCURACY, VALIDITY, AND RELIABILITY	
<i>Is there a process to validate non-standardized assessments?</i>	<ul style="list-style-type: none"> • Non-DCAS assessments <u>approved by the DDOE</u> that are standards-based and show fall-to-spring growth or proficiency • Teacher-selected pre- and post-assessments must meet criteria of <u>internal measures rubric</u>
<i>Are there procedures to ensure that observation scores are valid?</i>	<ul style="list-style-type: none"> • State development coaches spend four hours a week training administrators to ensure proper evaluation of teachers and administrators <ul style="list-style-type: none"> - The position of development coach was created through Delaware’s Race to the Top application to work with districts and administrators across the state to ensure DPAS II- Revised is being implemented correctly - Development coaches do not observe or rate teachers • District expert evaluators (district office administrators, lead mentors, or development coaches) will model effective evaluation and monitor observations and evaluations
<i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i>	No
<i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i>	No
USING AND REPORTING EVALUATION RESULTS	
<i>What are the rating categories?</i>	<p>Teachers will receive an overall rating of highly effective, effective, needs improvement, or ineffective.</p> <ul style="list-style-type: none"> • Highly effective: teacher earns satisfactory ratings on at least four of the five components, including rating of “exceeds” on growth component (component 5) • Effective: teacher earns satisfactory ratings on at least three of the five components, but does not receive a rating of “exceeds” on growth component (component 5) • Needs improvement: teacher earns one or two satisfactory ratings on five components, including growth component (component 5); OR, teacher earns three or four satisfactory ratings on five components, plus rating of “unsatisfactory” on growth component (component 5) • Ineffective: teacher earns zero to two satisfactory ratings on five components, plus rating of “unsatisfactory” on growth component (component 5); OR, teacher’s overall rating is needs improvement for three consecutive years
<i>How are results reported?</i>	<ul style="list-style-type: none"> • Individual teacher results are reported to teachers and administrators only • RTTT commitment is to report evaluation results publicly by district, school, and teacher preparation program in the near future

<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<p>Yes, evaluation results are or will be tied to pay, professional development, and licensure. Other consequences are delivered at the district's discretion.</p> <ul style="list-style-type: none"> • Professional development/improvement plan <ul style="list-style-type: none"> - A teacher will receive an improvement plan for any of the following situations: <ul style="list-style-type: none"> » Teacher does not meet expectations set by evaluators during conferences » Teacher receives "unsatisfactory" rating on a nonacademic component during a lesson observation » Teacher is rated unsatisfactory on any of the nonacademic components (Components 1–4) on overall evaluation » Teacher receives an overall rating of ineffective or needs improvement. Teacher receives a rating of unsatisfactory on growth component (Component 5) (will not apply during interim year 2011–12) • Pay <ul style="list-style-type: none"> - Beginning in 2012–13, RTTT attraction and retention bonuses given to teachers who maintain highly effective rating in a high-needs school or transfer to a high-needs school - In year one, only DCAS reading and math teachers will be eligible • Licensure <ul style="list-style-type: none"> - A teacher who began teaching in the 2010–11 school year must have at least two years of "satisfactory" ratings in growth component (Component 5) within a three-year period to secure her continuing license. The interim 2011–12 school year will not count as part of the three-year period. • Other <ul style="list-style-type: none"> - Teachers who excel in any criterion or element of evaluation, or new teachers who demonstrate substantial professional growth, are eligible for a "commendation," additional comments in the evaluation that recognize outstanding achievement - If there is a "pattern of ineffective teaching" (see p. 58 of Guide) a teacher may be dismissed at the local school board's discretion - Rewards such as promotions or increased responsibilities are delivered at district's discretion
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<ul style="list-style-type: none"> • Teacher may submit a written challenge to the evaluator's supervisor, an administrator, for: <ul style="list-style-type: none"> - Any lesson observation rating - Any component or overall rating on final evaluation • Teacher must submit specific information to evaluator's supervisor within 15 working days of teacher's receipt of evaluation • Supervisor must hold a challenge hearing with teacher to review and discuss challenge and evaluation record within 15 working days of receipt of challenge • Supervisor must issue a written decision within 15 working days of challenge hearing. If challenge is denied, decision must state reasons for denial. Supervisor's decision is final <p>Note: If evaluator and teacher work in the same building, supervisor will be a district or charter management organization-level credentialed evaluator.</p>
<p><i>If performance results are available, what is the score distribution?</i></p>	<p>Data are currently unavailable. However, Delaware anticipates a distribution for overall rating (for 2013–14) of:</p> <ul style="list-style-type: none"> • 30% Highly effective • 50% Effective • 15% Needs improvement • 5% Ineffective

SITE PROFILES

TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT	
<i>What role did teachers play in developing the evaluation system?</i>	<ul style="list-style-type: none"> • More than 500 teachers participated in working groups to develop internal measures of student growth. These working groups are ongoing <ul style="list-style-type: none"> - Teachers were nominated by districts and unions - Teachers make high-level decisions, which are reviewed by DDOE and the Technical Advisory Group • DDOE hosted forums to gather teacher input throughout design and revision of DPAS II-Revised
<i>What role did the union play in developing the evaluation system?</i>	<ul style="list-style-type: none"> • Union involved in drafting regulations and ensuring compliance with RTTT commitments • Union makes recommendations to recruit teachers and administrators to develop aspects of DPAS II-Revised • Former president of union is now special assistant for DPAS II-Revised in the DDOE
<i>Is there a formal process for revising the evaluation system over time?</i>	<ul style="list-style-type: none"> • Every year since the inception of DPAS II-Revised, DDOE has contracted with Progress Education Corporation to conduct a full <u>evaluation of the system</u>. Progress studies current manuals, forms, and historical documents; administers surveys; conducts interviews; and facilitates focus groups to gather qualitative and quantitative data on: <ul style="list-style-type: none"> - Criteria used in the DPAS II-Revised system - Forms used for evaluations - Manageability of the entire system - Accuracy and reliability of the data being used in the system - Usefulness of the training sessions and manuals - Needed modifications - Efficacy of the DPAS II-Revised program in achieving quality assurance and professional growth • Using data collected, Progress develops a set of recommendations for revising the system for the next year

GLOSSARY	
<i>Component</i>	<ul style="list-style-type: none"> • One of the five areas of teacher practice and responsibility that teachers are evaluated against. • Components 1–4 are nonacademic • Component 5 is student growth
<i>Criterion</i>	<ul style="list-style-type: none"> • Subcategory of knowledge/skills within a component
<i>Development coach</i>	<ul style="list-style-type: none"> • Development coaches help deliver regional trainings to all administrators and work with administrators to norm and calibrate observation results. They do not observe or rate teachers. • Development coaches are typically former administrators or veteran teacher leaders. • There are nine coaches working in 60 schools across the state. Their contract is for two years. Next year, 10 coaches will serve 80 schools. • The position of development coach was created through Delaware’s Race to the Top application to work with districts and administrators across the state to ensure DPAS II-Revised is being implemented correctly
<i>Experienced teacher</i>	<ul style="list-style-type: none"> • A teacher holding a valid and current Continuing or Advanced License issued prior to August 1, 2003³³ • All other teachers are considered “novice”
<i>Expert evaluator</i>	<ul style="list-style-type: none"> • District office administrators, lead mentors, or development coaches that model effective evaluation and monitor observations and evaluations
<i>External measures</i>	<ul style="list-style-type: none"> • Non-DCAS assessments <u>approved by the DDOE</u> that are standards-based and show fall-to-spring growth • Examples: district-wide Measures of Academic Progress (MAP assessments) or IEP progress
<i>Internal Measures</i>	<ul style="list-style-type: none"> • May be internal measures: teacher-selected pre- and post-assessments • Must be approved by DDOE and meet criteria of <u>internal measures rubric</u> • Measures are selected by teacher and evaluator between August and October of the school year
<i>New (or Novice) teacher</i>	<ul style="list-style-type: none"> • Teacher holding a valid and current Initial License
<i>Pattern of Ineffective Teaching</i>	<ul style="list-style-type: none"> • See the <u>Guide, p. 58</u> for chart that shows possible consecutive ratings considered to be a pattern of ineffective teaching

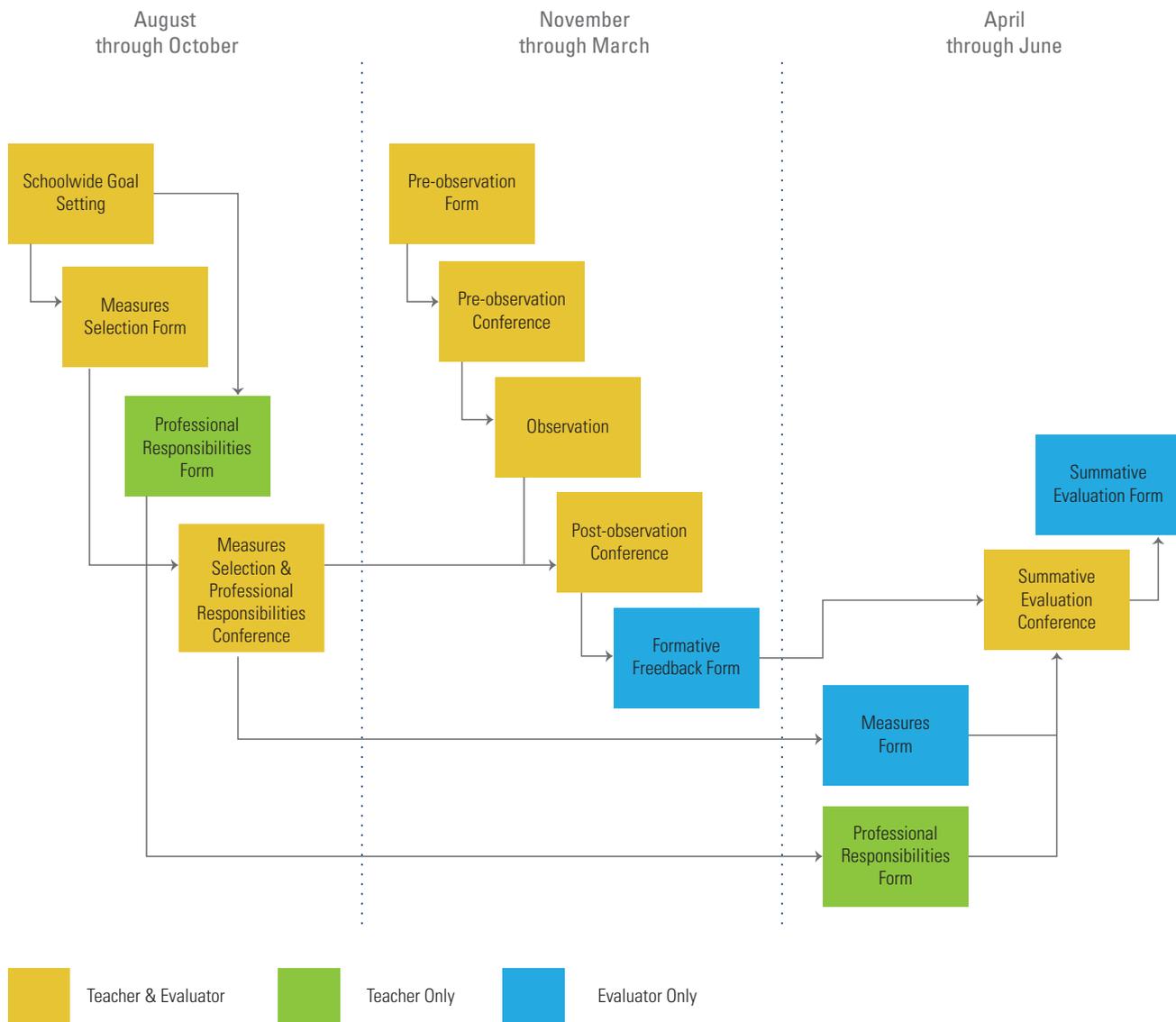
³³Delaware does not award “tenure,” so teachers are either considered “novice” or “experienced” (holding a valid and current Continuing or Advanced License).

SITE PROFILES

TIMELINE

DPAS II-REVISED - APPRAISAL CYCLE

This diagram illustrates a suggested timeline for the DPAS II evaluation process for teachers. This timeline is intended as a guide, not a mandate.



Note: Observation feedback process may be repeated using announced or unannounced observations

SOURCES

Delaware Department of Education. (2007). *Competitive sealed proposal for the evaluation of the Delaware Performance Appraisal System Second Edition (DPAS II)*. Retrieved May 1, 2012, from <http://www.doe.k12.de.us/rfp/DPASIIIRFP2008FINAL.pdf>

Delaware Department of Education. (2011). *DPAS II guide for teachers*. Retrieved November 20, 2011, from <http://www.doe.k12.de.us/csa/dpasii/ti/DPASII TeachFullGuide.pdf>

Delaware Department of Education. (2011). *DPAS II summative ratings for 2011–2012*. Retrieved November 20, 2011, from <http://www.doe.k12.de.us/csa/dpasii/summative2011final.shtml>

Delaware Department of Education. (2011). *Non-regulatory guidance for the DPAS II process for teachers and specialists*. Retrieved November 20, 2011, from <http://www.doe.k12.de.us/csa/dpasii/new/DPASII NonregGuidance-8-4-11.pdf>

Delaware Department of Education. "State Profile." Retrieved November 20, 2011 at: <http://profiles.doe.k12.de.us/SchoolProfiles/State/Account.aspx>

Delaware State Education Association. (2011). "Getting DPASII-REVISED right for 2011-2012." DSEA Action! Vol. 32, No. 1. Available: <http://digital.copcomm.com/issue/40928>

Progress Education Corporation. Annual DPAS II Evaluation Reports. Available: <http://www.doe.k12.de.us/csa/dpasii/default.shtml>

RHODE ISLAND

RHODE ISLAND MODEL EDUCATOR EVALUATION SYSTEM

OVERVIEW

Rhode Island began efforts to overhaul its teacher evaluation system in 2009. Those efforts received a boost in 2010 when the state won a Race to the Top grant, leading the state to roll out the Rhode Island Model Educator Evaluation System in 2011.

The Rhode Island model will measure student growth using two different measures: a Growth Model score (available 2012–13) and a Student Learning Objective score. The Growth Model score will measure a student’s growth on the state exam in comparison to his/her academic peers by assigning students a growth percentile based on their relative growth. In addition, all teachers work with their administrators to set student learning objectives (SLOs), a long-term academic goal that is specific and measureable. Teachers then receive a score based on the degree to which students reached the SLO. If a Growth Model score is not available for a particular teacher, the SLO score will count as the only student growth measure. Teachers also receive scores for nonacademic measures, including professional practice and professional responsibility. Teachers receive ratings for each component, which are then put into a matrix to calculate a final evaluation score.

The Rhode Island Department of Education (RIDE) has decided to implement a limited new evaluation system (fewer observations, no growth model rating, overall rating used for developmental purposes only) during the 2011–12 school year so district leaders can receive feedback and educators can have an adjustment period before tying consequences directly to ratings in 2012–13. RIDE faces the challenge of refining the system based on feedback from 2011–12 in time for full implementation in 2012–13. This timeline requires quickly incorporating numerous changes to the system by June 2012, when training begins for full implementation.

ABOUT THE SCHOOL SYSTEM

<i>Number of students</i>	143,793
<i>Number of schools</i>	300
<i>Number of teachers</i>	14,260

BACKGROUND

<i>Corresponding legislation</i>	No new legislation, but Board of Regents adopted into regulation new Educator Evaluation Standards in 2009
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	Yes, in Spring 2011
<i>In effect since</i>	Gradual implementation began in all districts in 2011–12, with fewer observations and no RI Growth Model rating. Overall ratings for teachers are calculated, but they are used for developmental purposes only (see Guide, p. 8 , for more details on gradual implementation requirements for 2011–12).

<p><i>Are there plans for additional phases/components?</i></p>	<ul style="list-style-type: none"> • Full implementation in 2012–13, with the exception of RI Growth Model rating; all teachers will receive an overall rating • Growth Model rating will be included in the overall rating in 2013–14 • Evaluation model for support professionals (social workers, nurses, counselors, school psychologists, and central office staff) in development
<p><i>Who gets evaluated?</i></p>	<p>All teachers and administrators</p>
<p><i>How often do teachers receive a rating?</i></p>	<p>Once a year</p>
<p><i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i></p>	<ul style="list-style-type: none"> • The state provides an evaluation model, but allows districts to propose other models that meet state requirements • Each district will have a District Evaluation Committee (DEC) that will make implementation decisions related to: <ul style="list-style-type: none"> - Teacher appeals - Improvement plans - Local assessments and rubrics
<p>STUDENT ACHIEVEMENT MEASURES</p>	
<p><i>Assessments to measure student achievement</i></p>	<p><u>New England Common Assessment Program (NECAP)</u></p> <ul style="list-style-type: none"> • Grades 3–8 • Reading and math
<p><i>Are any new assessments in development to use with teacher evaluations?</i></p>	<p><u>Partnership for Assessment of Readiness for College and Careers (PARCC)</u></p> <ul style="list-style-type: none"> • PARCC will provide K–12 assessments aligned to Common Core State Standards in English language arts and math • Assessments throughout the year will include summative and non-summative tests, a speaking and listening component, and variety of constructed response items, performance-based tasks, and computer-scored items • PARCC assessments will begin in 2014–15 • Once enough data are available, RIDE will use PARCC results to measure proficiency and growth for grades 3–11

SITE PROFILES

<p><i>What academic/achievement criteria are included?</i></p>	<p>The student learning component of the evaluation is based on two criteria: a teacher's Growth Model score (available 2013–14 and her Student Learning Objective score:</p> <ul style="list-style-type: none"> • The <u>Rhode Island Growth Model</u> (RIGM) score measures student growth on the state exam in comparison to their academic peers by assigning students a growth score. <ul style="list-style-type: none"> - Teachers in tested grades and subjects: The RIGM score is based on student growth on the NECAP reading and math tests for grades 3–7 - Contributing educators <ul style="list-style-type: none"> » All teachers who contribute to the literacy or mathematics development of students (e.g., special educators, co-teaching educators, specialists) are considered “contributing educators” » RIDE provides a <u>policy guide</u> to districts that offers guidelines for identifying contributing educators and suggestions for how growth should be attributed to them - All other teachers. Teachers who do not contribute to the literacy or mathematics development of students in tested grades and subjects do not receive an RIGM score - Growth score. RIGM uses median growth percentiles, rather than a value-added measure³⁴ <ul style="list-style-type: none"> » Each student's annual growth on his NECAP score is compared to the growth of his or her academic peers (students with a similar test score history) » Each student then receives a Student Growth Score (SGS) based on his relative performance » Teachers receive a score of Low Growth, Typical Growth, or High Growth based on the median SGS of her class over two years (cut scores for each level have not yet been determined) • <u>Student learning objectives</u> (SLOs) are specific and measurable goals aligned to state standards and based on available prior student data <ul style="list-style-type: none"> - All teachers work with evaluators and administrators to set two to four SLOs. Objectives must be <u>approved</u> based on three criteria (see <u>student learning guide, p. 17</u>): <ul style="list-style-type: none"> » Alignment to state standards » Rigor of target » Quality of evidence to measure progress - SLOs may measure either growth or mastery (See more examples <u>here</u>) <ul style="list-style-type: none"> » Example of mastery objective: <i>Students will demonstrate an understanding of United States government (local, state, national) by identifying elected leadership titles/basic role at different levels of government (e.g., mayor is the leader of a city, governor is the leader of the state, president is the leader of the country)</i> » Example of growth objective: <i>Students will improve their ability to compose informative/explanatory writing, including a clear thesis statement, organized body paragraphs with supporting evidence, and a concluding statement</i> - Teachers and evaluators will meet midyear to discuss progress toward their SLOs and revise them as needed - At the end of the year, evaluators will review results of assessments and other evidence to determine the extent to which the SLOs were met (<u>see rubric on student learning guide, p. 21</u>)
--	--

³⁴See the introductory brief for more information on student growth percentiles and value-added measures.

<p><i>How much do the student achievement measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • Student Learning score is calculated using the Student Learning Matrix (Guide, p. 65), and includes: <ul style="list-style-type: none"> - Results from RI Growth Model for tested subjects and grades and “contributing educators” where defined (to be included in 2012–13) - Attainment of student learning objectives • Teachers of untested grades and subjects will receive a Student Learning score based solely on student learning objectives • For 2011–12, the Student Learning score will only include attainment of student learning objectives for all teachers • The final teacher rating is calculated using a matrix that combines the Student Learning score and Professional Practice and Professional Responsibilities score (Guide, p. 66) <ul style="list-style-type: none"> - Because the final ratings are calculated through a matrix, the weight of each component varies depending on the teacher’s performance on other pieces, but student performance accounts for roughly half of a teacher’s final evaluation score
<p>NONACADEMIC MEASURES</p>	
<p><i>What nonacademic criteria are included?</i></p>	<p>The Rhode Island model uses two nonacademic components: professional practice and professional responsibilities.</p> <ul style="list-style-type: none"> • Professional Practice criteria are planning and preparation; classroom instruction; classroom environment; and assessment, reflection, and improvement • Professional Responsibilities criteria are collaborating and contributing to the school community; belief in and advocating for students; creating a culture of respect; and exercising professional judgment and development
<p><i>How much do the nonacademic measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • Professional Practice and Professional Responsibilities score is calculated using a PPxPR matrix (Guide, p. 64) • The final teacher rating is calculated using a matrix that combines the Student Learning score and Professional Practice and Professional Responsibilities score (Guide, p. 66) <ul style="list-style-type: none"> - Because the final ratings are calculated through a matrix, the weight of each component varies depending on the teacher’s performance on other pieces, but professional practice and responsibilities count for roughly half of a teacher’s final evaluation score
<p>CLASSROOM OBSERVATIONS</p>	
<p><i>Who are the observers who rate teachers' effectiveness?</i></p>	<p>There are two kinds of evaluators in the Rhode Island Model: primary and complementary.</p> <ul style="list-style-type: none"> • Primary evaluators: <ul style="list-style-type: none"> - Usually principals or assistant principals, but LEAs may determine who primary evaluators will be, as long as they receive training - Primary evaluators have sole responsibility for assigning evaluation ratings • Complementary evaluators: <ul style="list-style-type: none"> - May be a department head, content-area expert, or administrator from within the school who assists primary evaluator by conducting observations, gathering evidence, or providing feedback - May be an Intermediary Service Provider (ISP) that RIDE assigns to districts to train and guide principals during the first few years of implementation - ISPs are usually retired administrators, consultants, or district officials - There are approximately 25 ISPs working across all of the state’s 52 school districts

SITE PROFILES

<p><i>What training do observers receive?</i></p>	<p>All evaluators (primary and complementary) must attend a four-day academy of training sessions led by experienced teachers and administrators trained by RIDE.</p> <ul style="list-style-type: none"> • Each preliminary session focuses on one aspect of the evaluation system: <ul style="list-style-type: none"> - Student learning - Professional Growth Plans - Observations and feedback - Conferencing • Evaluators will also receive follow-up training provided by RIDE throughout the school year • Beginning in 2012–13, RIDE will set expectations that all evaluators must meet in order to successfully complete training
<p><i>Are observers compensated for this work?</i></p>	<p>ISPs are compensated for their work. Over time, the ISP position will be transitioned out as Race To The Top funding expires</p>
<p><i>How often are teachers observed?</i></p>	<p>Teachers receive at least four observations each year, including:</p> <ul style="list-style-type: none"> • At least one formal observation: at least 30 minutes • Informal observations: about 15 minutes
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<ul style="list-style-type: none"> • Formal observations: at least one must be conducted before midyear conference • Informal observations: no time frame
<p><i>Are observations announced or unannounced?</i></p>	<ul style="list-style-type: none"> • Formal observations: announced • Informal observations: unannounced
<p><i>What is the basis for the observation rubric?</i></p>	<p>RIDE reviewed TAP, Danielson, CLASS, and IMPACT rubrics while designing the system's observation rubric</p>
<p><i>What is included in the observation scoring rubric?</i></p>	<p>Observation Rubric has criteria related to:</p> <ul style="list-style-type: none"> • Planning and Preparation • Classroom Instruction • Classroom Environment • Assessment, Reflection and Improvement • Professional Responsibilities
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • RIDE is monitoring the implementation of ratings during its gradual implementation year • Additional calibration tools will be introduced during the 2012–13 implementation year
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • All observations followed by written feedback to teachers within two to three school days • Formal observations require post-observation conferences

ACCURACY, VALIDITY, AND RELIABILITY

<p><i>Is there a process to validate non-standardized assessments?</i></p>	<p>All assessments used to measure student learning objectives must be approved by evaluators using RIDE criteria (Guide, p. 50).</p>
<p><i>Are there procedures to ensure that observation scores are valid?</i></p>	<p>See below</p>
<p><i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i></p>	<ul style="list-style-type: none"> • At the state level, RIDE will: <ul style="list-style-type: none"> - Periodically audit the evaluation process to ensure that evaluations are fair and accurate, and that they adhere to the Rhode Island Educator Evaluation System Standards • At the district level, District Evaluation Committees are responsible for: <ul style="list-style-type: none"> - Reviewing the accuracy and utility of the data produced - Reviewing the evaluation decisions made for fairness and consistency - Providing procedural safeguards to ensure the integrity of the system, including evaluation appeals - In the event that an evaluation process yields a contradictory outcome (e.g., a teacher has an extremely high Student Learning rating and an extremely low rating in Professional Practice and Professional Responsibilities), a review of the evaluation will be conducted at the district level • RIDE has committed to using RTTT funding to create an online Educator Performance Support System (EPSS) that maintains all evaluation data that RIDE can monitor
<p><i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i></p>	<p>Yes, see above.</p>

SITE PROFILES

USING AND REPORTING EVALUATION RESULTS	
<i>What are the rating categories?</i>	Teachers receive a summative rating of highly effective, effective, developing, or ineffective.
<i>How are results reported?</i>	<ul style="list-style-type: none"> • Individual teacher results are reported to teachers and administrators only • Teachers and administrators will have access to evaluation results throughout the year on RIDE’s computer-based Educator Performance and Support System (EPSS) beginning in 2012–13
<i>Does the system allow teacher evaluation results to be reported by the training programs teachers attended?</i>	<ul style="list-style-type: none"> • RIDE will publish report cards for all educator preparation programs in the state based on teacher evaluation results • Report cards will be available to the public at the end of 2014–15 school year
<i>Are there consequences tied to evaluation results? If yes, what are they?</i>	<p>Yes, evaluation results are tied to promotion, professional development, and certification. RIDE plans to tie student assignment to evaluation results in the near future.</p> <ul style="list-style-type: none"> • Professional development/improvement plan <ul style="list-style-type: none"> - All teachers are required to have an annual professional development plan - Teachers who are rated developing or ineffective will be placed on an improvement plan for the following year • Promotion (effective January 2012) <ul style="list-style-type: none"> - RIDE issues differentiated teacher certificates (initial, professional, and advanced) that recognize different stages of development and accomplishment for teachers. Districts are encouraged to create career ladders for teachers. See regulations here • Student assignment (in development) <ul style="list-style-type: none"> - RI has <u>committed</u> to using RTTT funding to develop a system that ensures that a child will not have an ineffective teacher two years in a row • Certification (regulation effective January 2012) <ul style="list-style-type: none"> - Rhode Island educators who are rated highly effective, effective or developing on annual evaluations will be eligible for renewal of their certification - Experienced educators who receive evaluations of ineffective for five years in a row will not be able to renew their certification - New teachers with initial certification who receive evaluations of ineffective for three years in a row will lose their certification - RIDE is developing a process of training and supervised student teaching for teachers to be reinstated if they lose their certification - Certification is required to teach in all Rhode Island public schools • Performance pay, increased responsibility, tenure, and dismissal consequences are district-level decisions <ul style="list-style-type: none"> - If a teacher is ineffective for two years, RIDE recommends that districts move the teacher toward dismissal
<i>Can teachers appeal their ratings? If yes, how?</i>	Appeals will be handled at the district level in accordance with district policy and practice, collective bargaining agreements, and/or processes set by the District Evaluation Committee. State guidelines to govern district-level appeals are under development.
<i>If performance results are available, what is the score distribution?</i>	Data are unavailable.

TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT

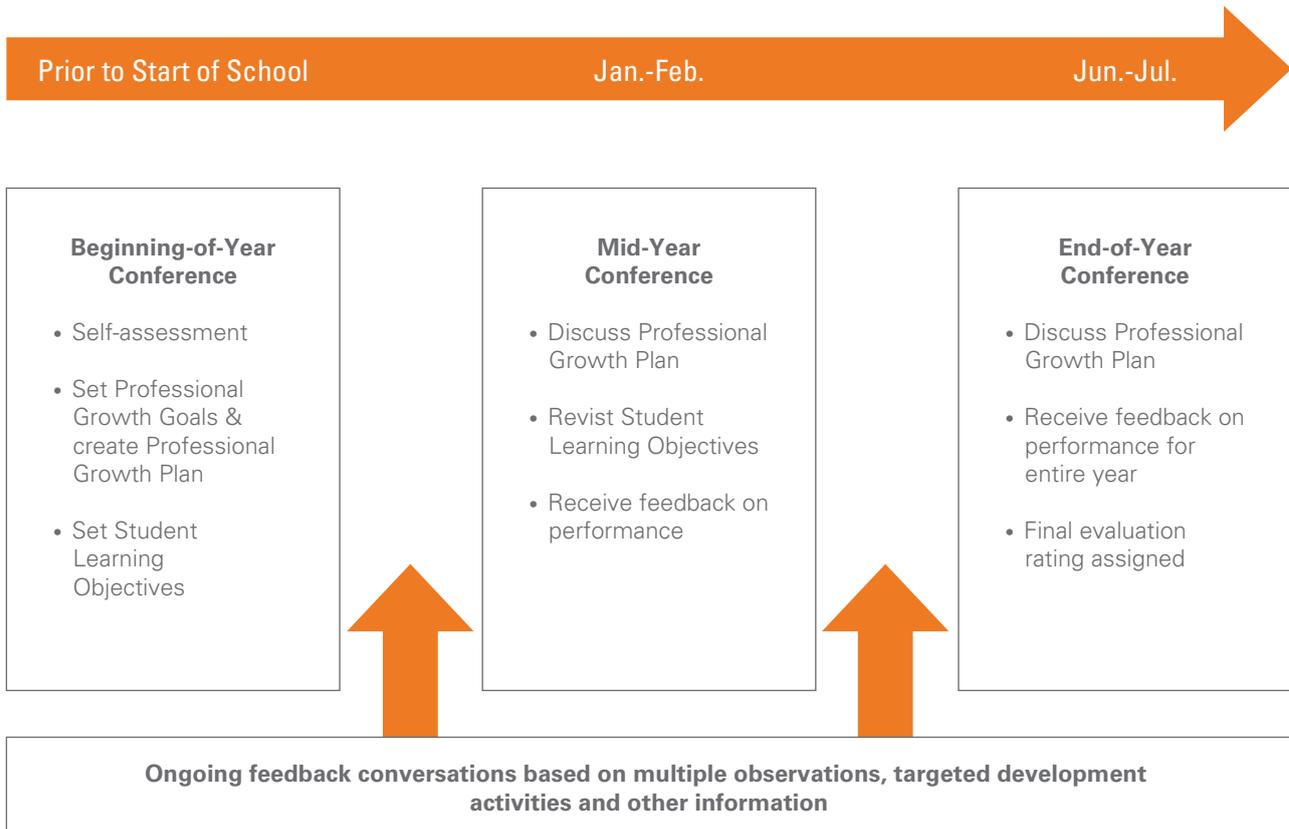
<p><i>What role did teachers play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • During development of the evaluation system, RIDE created the Advisory Committee for Educator Evaluation Systems (ACEES) that <ul style="list-style-type: none"> - Was composed of parents, students, and educators - Advised RIDE on design of the RI Model - Worked alongside the Technical Advisory Committee • Teachers also sit on District Evaluation Committees, which make implementation decisions at the district level and provide feedback to RIDE
<p><i>What role did the union play in developing the evaluation system?</i></p>	<p>Union members sit on District Evaluation Committees, which make implementation decisions at the district level and provide feedback to RIDE.</p>
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<p>The model will continue to be improved based on:</p> <ul style="list-style-type: none"> • Feedback from District Evaluation Committees, composed of union leaders, school committee members, parents, teachers and administrators • Feedback from the state Technical Advisory Committee (TAC), a group of national education and assessment experts who advised on methodologies and implementation • Formal reviews of evaluation data

SITE PROFILES

GLOSSARY	
<i>Complementary evaluator</i>	<ul style="list-style-type: none"> • May be a department head, content-area expert, or administrator from within the school who assists primary evaluator by conducting observations, gathering evidence, or providing feedback • May be an Intermediary Service Provider (ISP) whom RIDE assigns to districts to train and guide principals during the first few years of implementation; ISPs are usually retired administrators, consultants, or district officials
<i>District Evaluation Committee (DEC)</i>	<ul style="list-style-type: none"> • Oversees the implementation of educator evaluation in each local school system and ensures that the system is fairly and accurately administered
<i>Educator Performance Support System (EPSS)</i>	<ul style="list-style-type: none"> • RIDE’s online database that maintains all evaluation data • Teachers and administrators will have access to evaluation results throughout the year on EPSS • RIDE will use EPSS to monitor data for distribution and validity • EPSS will be implemented in 2012–13
<i>Intermediary Service Provider (ISP)</i>	<ul style="list-style-type: none"> • RIDE assigns ISPs to districts to train and guide principals during the first few years of implementation • ISPs are usually retired administrators, consultants, or district officials • There are approximately 25 ISPs working across all of the state’s 52 school districts • ISPs are compensated for their work • Over time, the ISP position will be transitioned out as Race To The Top funding expires
<i>Primary evaluator</i>	<ul style="list-style-type: none"> • May be principals, assistant principals, or department heads • Have sole responsibility for assigning evaluation ratings
<i>Student learning objectives</i>	<ul style="list-style-type: none"> • Learning goals set by teachers and approved by evaluators • Must be specific and measurable, aligned to state standards, and based on available prior student data • Objectives may measure either growth or mastery (see student learning guide for more information)
<i>Technical Advisory Committee (TAC)</i>	<ul style="list-style-type: none"> • Committee composed of national experts on assessment, performance management, and evaluation systems • Advises RIDE on all technical aspects of the model, including rating methodologies, Student Learning Objectives and the Rhode Island Growth Model

TIMELINE

TEACHER EVALUATION AND DEVELOPMENT PROCESS



SOURCES

Rhode Island Department of Education. *Office of Educator Quality and Certification*. <http://www.ride.ri.gov/educatorquality/educatorevaluation/Default.aspx>

Rhode Island Department of Education. *Rhode Island Growth Model*. <http://www.ride.ri.gov/assessment/RIGM.aspx>

Rhode Island Department of Education. (2011). *The Rhode Island Model: Guide to evaluating building administrators and teachers, 2011–2012*. Retrieved December 2, 2011, from <http://www.ride.ri.gov/educatorquality/educatorevaluation/Docs/RIModelGuide.pdf>

Rhode Island Department of Education. (2011). *Total populations of students, teachers and schools in Rhode Island*. Retrieved December 13, 2011, from <http://infoworks.ride.ri.gov/ri/statewide-data>

TENNESSEE

TENNESSEE EDUCATOR ACCELERATION MODEL (TEAM)

OVERVIEW

Tennessee was one of just two states to receive a grant in the first round of the federal Race to the Top competition in 2010. As part of its application, Tennessee outlined a plan for a new statewide teacher evaluation model, which led to the Tennessee Educator Acceleration Model (TEAM).

Half of teachers' evaluation scores are based on nonacademic measures (observations), while the other half are academic (student growth and student achievement). Every teacher in the state is observed by a school administrator or coach at least four times a year using a modified version of the Teacher Advancement Program (TAP) rubric. Tennessee relies on its value-added model to measure student growth. TEAM uses individual value-added measures for teachers of tested grades and subjects, and a school-level value-added measure for teachers who do not teach a tested grade or subject. In addition, teachers select a second student achievement measure to include in their evaluation results from a pre-approved list. Additional measures are being developed.

Although districts are responsible for making many decisions related to the implementation of TEAM—including who can evaluate teachers and how to sequence observations—several aspects are mandated in state law. Most notably, all teacher evaluation systems must include a measure of teacher practice, student growth, and student achievement. State law also mandates how much each component counts in a teacher's final score, and requires that annual evaluations be a factor in personnel decisions, including promotion, retention, tenure, and compensation.

Because districts retain a lot of ownership over TEAM, one of the state's greatest challenges is balancing its role in implementation with that of districts. The state is also thinking through ways to improve stakeholder buy-in locally, including accelerating the pace of change at the district and school levels.

All forms used for TEAM can be found [here](#).

ABOUT THE SCHOOL SYSTEM

<i>Number of schools</i>	1,736
<i>Number of students</i>	933,703
<i>Number of teachers</i>	64,229

BACKGROUND

<i>Corresponding legislation</i>	First to the Top Act
<i>Year legislation passed</i>	2010
<i>Was the system piloted?</i>	<ul style="list-style-type: none"> • Yes, the qualitative measures were piloted in a subset of schools in spring 2011 • Tennessee has generated value-added scores for nearly two decades
<i>In effect since</i>	2011–12 (full implementation, as mandated in the state law)

<p><i>Who gets evaluated?</i></p>	<p>Teachers and principals</p>
<p><i>How often do they receive a rating?</i></p>	<p>Once a year</p>
<p><i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i></p>	<ul style="list-style-type: none"> • Yes, Tennessee provides a state-created evaluation model, but enables districts to propose other models that meet state requirements. Fourteen districts currently use one of three alternative models • All districts must use a system that includes the growth and student achievement components described above, but districts may implement their own observation systems if they meet state board approval
<p>STUDENT ACHIEVEMENT MEASURES</p>	
<p><i>Assessments to measure student achievement</i></p>	<ul style="list-style-type: none"> • The state test is the <u>Tennessee Comprehensive Assessment Program (TCAP)</u>. It includes: <ul style="list-style-type: none"> - Achievement Assessment <ul style="list-style-type: none"> » Grades: 3–8 » Subjects: reading, language arts, mathematics, science, and social studies - TCAP writing assessment <ul style="list-style-type: none"> » Grades 5, 8, and 11 - TCAP Secondary Assessment Tests <ul style="list-style-type: none"> » Grades: 9–12, as appropriate » Subjects: English 1, English II, English III, algebra I, geometry, algebra II, U.S. history, biology I - The TCAP includes several alternative assessments for students with disabilities • <u>End-of-course assessments (EOCs)</u> for some high school classes (English 1, English II, English III, algebra I, geometry, algebra II, U.S. history, biology I)
<p><i>Are any new assessments in development to use with teacher evaluations?</i></p>	<ul style="list-style-type: none"> • The state is working with teams of educators to identify and/or develop alternative measures of student growth • The state must approve any new measures to ensure they meet the appropriate criteria • Once approved, all LEAs would have the option to use these approved measures, but could still choose to use school-level scores for some or all educator groups outside of TCAP/EOC teachers
<p><i>What academic/achievement criteria are included?</i></p>	<p>There are two academic criteria: student growth data and student achievement.</p> <ul style="list-style-type: none"> • Student growth data. Tennessee has been using the Tennessee Value-Added Assessment System (TVAAS) for 18 years, which calculates value-added scores for individual teachers using TCAP and EOC data <ul style="list-style-type: none"> - For teachers with individual value-added data (approximately 45% of all teachers) <ul style="list-style-type: none"> » The state uses their individual value-added data - For teachers without individual value-added data (approximately 55% of all teachers) <ul style="list-style-type: none"> » The state has provided <u>guidance on appropriate value-added scores</u>, which measure growth at the school level using data from all tested areas, literacy only, numeracy only, or literacy and numeracy only <ul style="list-style-type: none"> » E.g., library media specialist—overall, literacy » E.g., early grades—overall or math/literacy (from feeder schools)

SITE PROFILES

<p><i>What academic/achievement criteria are included (cont.)?</i></p>	<ul style="list-style-type: none"> • <u>Student achievement data.</u> - Selected by the teacher and supervisor from a <u>list of state-approved board options</u> (see p. 3) - Examples include: other national or state-used assessments, postsecondary matriculation, ninth-grade retention/promotion, graduation rates, completion/success of advanced coursework, etc. - Educators and evaluators work together to agree upon appropriate achievement measures that are aligned as closely as possible with the teacher’s primary responsibilities and are measurable using student performance data
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<p>Student academic performance counts for 50% of a teacher’s final rating:</p> <ul style="list-style-type: none"> • Student growth—35% • Student achievement data—15% • Teachers in the top three quintiles of student growth can choose for their student growth to count for 50% in lieu of selecting another achievement measure
<h3>NONACADEMIC MEASURES</h3>	
<p><i>What nonacademic criteria are included?</i></p>	<p>Half of a teacher’s evaluation score is based on observation in the following areas, based on the rubric for the Teacher Advancement Program (TAP):</p> <ul style="list-style-type: none"> • Planning • Environment • Professionalism • Instruction <p>(See a rubric for general education teachers here).</p>
<p><i>How much do the nonacademic criteria count in a teacher’s final rating?</i></p>	<p>The nonacademic criteria count for 50% of a teacher’s final evaluation score.</p>
<h3>CLASSROOM OBSERVATIONS</h3>	
<p><i>Who are the observers who rate teachers’ effectiveness?</i></p>	<p>Observers can include:</p> <ul style="list-style-type: none"> • Principals • Assistant principals • Other instructional leaders trained in the observation protocol (e.g., district-level supervisor, instructional coach, department or grade-level chairs, retired educators) • Districts are responsible for defining observers
<p><i>What training do observers receive?</i></p>	<ul style="list-style-type: none"> • TDOE provided a four-day training session for a team from every Tennessee school, including at least the principal and assistant principal for 2011–12 • Observers complete a certification test through the portal • State will offer refresher courses and support throughout the year as well • Trainers include TAP staff and contracted trainers
<p><i>Are observers compensated for this work?</i></p>	<p>No</p>

<p><i>How often are teachers observed?</i></p>	<p>Each teacher is observed at least four times a year, with new teachers receiving more observations. For example:</p> <ul style="list-style-type: none"> • Apprentice teachers: at least 6 times per year <ul style="list-style-type: none"> - 3 15-minute observations (covering 3 indicators from the rubric) - 3 lesson-length observations (covering 12 indicators from the rubric) - Apprentice teachers have completed an approved teacher preparation program, but do not yet have a professional license. Apprentice teachers can earn a professional license after they teach for three years in Tennessee and the school system submits evidence of a positive local evaluation. • Non-apprentice teachers: at least 4 (2 each semester) <ul style="list-style-type: none"> - 2 15-minute observations (covering three indicators from the rubric) - 2 lesson-length observations (covering 12 indicators from the rubric)
<p><i>Is there a timeframe within which observations must be conducted?</i></p>	<p>Half of all observations must take place in the first semester, while the other half must take place in the second.</p>
<p><i>Are observations announced or unannounced?</i></p>	<p>At least half of all observations will be unannounced.</p>
<p><i>What is the basis for the observation rubric?</i></p>	<p>Tennessee uses a modified version of the TAP rubric</p>
<p><i>What is included in the observation scoring rubric?</i></p>	<ul style="list-style-type: none"> • The modified version of the TAP rubric includes 19 indicators of effective teaching across three areas: planning, environment, and instruction • The rubric differs slightly for alternative educators and library/media specialists • Districts may choose to apply for a waiver to use other approved models
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • Trainers must pass a certification test administered online before they can conduct evaluations • The Tennessee Department of Education (TDOE) has said that it will analyze performance results by district and schools throughout the year, and where scores are not differentiated, TDOE will intervene with re-training
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • Teachers must receive written feedback from observation visits within one week • Observers must hold post-observation conferences with teachers
<p>ACCURACY, VALIDITY, AND RELIABILITY</p>	
<p><i>Is there a process to validate non-standardized assessments?</i></p>	<p>Tennessee regularly shares data reports with districts and engages in conversations with districts where the data indicate reason for concern.</p>
<p><i>Are there procedures to ensure that observation scores are valid?</i></p>	<ul style="list-style-type: none"> • Trainers must pass a certification test administered online before they can conduct evaluations • The Tennessee Department of Education (TDOE) has said that it will analyze performance results by district and schools throughout the year, and where scores are not differentiated, TDOE will intervene with re-training, coaching, and co-observation

SITE PROFILES

<p><i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i></p>	<ul style="list-style-type: none"> • There is not a formal process, but Tennessee Department of Education (TDOE) plans to analyze the data and follow up where there are inconsistencies • TDOE also has the authority to revoke alternative models if scores are not aligned with student test scores (see November 2011 recommendations from the State Board of Education) • Since student outcomes and observations are different types of measures, however, TDOE does not anticipate perfect correlation • TDOE will work with a research partner at Vanderbilt University to do an extensive review of Year 1 data in addition to the department's own analysis
<p><i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i></p>	<p>Not applicable. Observation scores are the only nonacademic measure, except in Memphis, where the district is piloting student surveys as a “Measures of Effective Teaching” site through the Bill & Melinda Gates Foundation.</p>
<h3>USING AND REPORTING EVALUATION RESULTS</h3>	
<p><i>What are the rating categories?</i></p>	<p>There are five rating categories:</p> <ul style="list-style-type: none"> • (5) Significantly above expectations • (4) Above expectations • (3) Meets expectations • (2) Below expectations • (1) Significantly below expectations
<p><i>How are results reported?</i></p>	<p>Individual teacher results are reported to teachers and administrators only</p>
<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<p>The First to the Top Act requires that annual evaluations be a factor in personnel decisions, including promotion, retention, tenure, and compensation. The details of such decisions, however, are left to the district except with respect to tenure.</p> <ul style="list-style-type: none"> • Tenure <ul style="list-style-type: none"> - Teachers are eligible for tenure after teaching in the same LEA for at least five years and attaining a rating in the top two evaluation categories in each of the past two years - Change does not apply to teachers who earned tenure before July 1, 2011
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<ul style="list-style-type: none"> • Teachers may not appeal their ratings • Teachers may <u>grieve</u> the fidelity of the evaluation process and the accuracy of the data (the correctness of the data associated with a teacher). See state policy for more details.
<p><i>If performance results are available, what is the score distribution?</i></p>	<p>2011–12 is the first year TEAM is being implemented, so scores are not yet available. Historically, the distribution of TVAAS scores is approximately:</p> <ul style="list-style-type: none"> • 5 (highest): 5–10% • 4: 15–25% • 3: 25–30% • 2: 15–25% • 1 (lowest): 2–3%

TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT

<p><i>What role did teachers play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • The governor appointed four public school teachers and two public school principals as part of the 15-member <u>Teacher Evaluation Advisory Committee (TEAC)</u> • Groups of 8 to 12 educators in each of the non-tested subject and grade groupings met to recommend growth measures • In the first year of implementation, Tennessee has received feedback on TEAM from more than 7,000 teachers through focus groups, Q&A sessions, and similar forums • Tennessee has administered a formal survey soliciting feedback through Vanderbilt University
<p><i>What role did the union play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • Throughout the Race to the Top application process and the development of TEAM, the state engaged in weekly phone calls with union representatives • Union leaders were represented on the TEAC
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<ul style="list-style-type: none"> • TDOE has said it is committed to studying the results to identify needed changes and refining the TEAM each summer • TDOE has already made one adjustment in response to feedback from school leaders. As of November, principals can conduct back-to-back observations followed by one conference (rather than two), requiring less time from principals

GLOSSARY OF TERMS

<p><i>Apprentice teacher</i></p>	<ul style="list-style-type: none"> • Apprentice teachers have completed an approved teacher preparation program, but do not yet have a professional license • Apprentice teachers can earn a professional license after they teach for three years in Tennessee and the school system submits evidence of a positive local evaluation
<p><i>Teacher Evaluation Advisory Committee (TEAC)</i></p>	<ul style="list-style-type: none"> • 15-member committee created by the General Assembly as part of the Race to the Top Act to make recommendations to the State Board of Education regarding guidelines and criteria for TEAM
<p><i>Tennessee Value Added Assessment System (TVAAS)</i></p>	<ul style="list-style-type: none"> • Tennessee’s value-added model, which draws on TCAP data

TIMELINE

Not available.

SITE PROFILES

SOURCES

Tennessee Department of Education. *About Tennessee First to the Top*. Retrieved from <http://www.tn.gov/firsttothetop/about.html>

Tennessee Department of Education. *Educators overview: New teacher and principal evaluation*. Retrieved from http://www.tn.gov/firsttothetop/docs/Educators_Overview.pdf

Tennessee Department of Education. *Teacher model*. Retrieved from <http://team-tn.org/teacher-model>

Tennessee State Board of Education. (2011). *Teacher and principal evaluation policy*. Retrieved from http://www.tn.gov/firsttothetop/docs/IV_C_Teacher_and_Principal_Evaluation_Policy.pdf

HILLSBOROUGH COUNTY, FL

EMPOWERING EFFECTIVE TEACHERS

OVERVIEW

Hillsborough County enrolls approximately 200,000 students across 250 schools in and around Tampa. In 2009, Hillsborough County Public Schools (HCPS) applied for and won a \$100 million grant from the Bill & Melinda Gates Foundation aimed at transforming how teachers are recruited, developed, rewarded, and retained. As one of Gates' four intensive partnerships for effective teaching nationwide, HCPS implemented a new teacher evaluation system, known as Empowering Effective Teachers (EET), in the 2010–11 school year.

HCPS's teacher evaluation system includes two main components: teacher practice (60%) and student achievement (40%). All teachers receive numerical scores for teacher practice based on observations. Teachers receive two to seven formal observations a year depending on experience and past effectiveness. In addition to school administrators, mentor teachers (for new teachers) and peer teachers (for more experienced teachers) also conduct observations. Both mentor and peer teachers are released from their teaching duties to fill these full-time positions dedicated to observing, holding conferences with, and evaluating teachers. The first time HCPS hired peer and mentor teachers, it received more than 600 applications for 116 positions.

Student achievement is measured using a teacher-level value-added score. HCPS has approved or developed at least one test for every grade and subject taught in its schools, although its value-added model uses performance data from as many tests as possible for each teacher. When enough data are available, the student performance portion of the teacher evaluation system will be based upon a three-year average to provide a more stable measure of student learning growth.

Currently, the Gates grant covers the approximately \$12 million per year that EET costs the district to train its evaluators, calibrate results, and pay the salaries for evaluator and mentor teachers. Hillsborough is committed to developing a financially sustainable plan beyond the life of the grant, although the district acknowledges that doing so poses a challenge.

ABOUT THE SCHOOL SYSTEM

<i>Number of students</i>	192,547
<i>Number of schools</i>	250
<i>Number of teachers</i>	12,468

BACKGROUND

<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	No
<i>In effect since</i>	2010–11
<i>Are there plans for additional phases / components?</i>	<ul style="list-style-type: none"> • 2011–12: HCPS will be updating its end-of-course exams, pre- and post-tests, and formative assessments • 2013–14: The Career Ladder Compensation Plan will go into effect (see slide 6)

SITE PROFILES

<p><i>Who gets evaluated?</i></p>	<ul style="list-style-type: none"> • Classroom teachers • Guidance counselors • Teachers of hospitalized/homebound students • Media specialists • Technology resource teachers • In 2012–13, HCPS plans to roll out new evaluations for assistant principals, social workers, and psychologists • New <u>principal evaluation</u> based on similar principles as teacher evaluation system was introduced in 2010–11
<p><i>How often do teachers receive a rating?</i></p>	<ul style="list-style-type: none"> • Tenured teacher: once a year (spring only) • Non-tenured teachers: twice a year (fall and spring), but fall score is formative
<p><i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i></p>	<p>Not applicable</p>
<p>STUDENT ACHIEVEMENT MEASURES</p>	
<p><i>Assessments to measure student achievement</i></p>	<ul style="list-style-type: none"> • Florida’s state assessment is the <u>Florida Comprehensive Assessment Test 2.0 (FCAT 2.0)</u> <ul style="list-style-type: none"> - Reading: grades 3–10 - Math: grades 3–8 - Science: grades 5 and 8 - Writing: grades 4, 8, and 10 (starting in 2013) • <u>End-of-course assessments</u> are computer-based tests designed to measure student achievement for specific middle school and high school courses <ul style="list-style-type: none"> - Algebra 1 - Biology 1 (starting May 2012) - Geometry (starting May 2012) - U.S. History (starting May 2013) - Civics (starting May 2014) • Pre-approved assessments <ul style="list-style-type: none"> - E.g. <u>Stanford 10</u> for math in grades 1 and 2, <u>Brigance Early Childhood Developmental Inventory</u> and <u>Batelle Developmental Inventory</u> for low-performing exceptional student education (ESE) students • District-created assessments • PSAT • AP exams • IB exams
<p><i>Are any new assessments in development to use with teacher evaluations?</i></p>	<p>A new district assessment team of eight coordinators and three supervisors is working with teachers to update the district’s end-of-course tests, pre- and post-tests, and formative assessments.</p>
<p><i>What academic/achievement criteria are included?</i></p>	<p>HCPS worked with the <u>University of Wisconsin’s Value-Added Research Center</u> to develop the <u>value-added measure</u> it uses for all teacher evaluation</p> <ul style="list-style-type: none"> • Produces a value-added score for teachers that will eventually be based on three years of student performance data • Includes student results from all available tests for a particular teacher
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<ul style="list-style-type: none"> • 40% is based on student learning gains • Note: By the third year of the program (and a teacher’s third year of teaching), the student learning growth portion of a teacher’s evaluation will be based upon a three-year average to provide a more stable measure of student learning growth

NONACADEMIC MEASURES	
<i>What nonacademic criteria are included?</i>	Principal and mentor/peer teacher evaluations
<i>How much do the nonacademic measures count in a teacher's final rating?</i>	60% is the teacher's written evaluation, which includes: <ul style="list-style-type: none"> • 30% determined by principal • 30% by mentor/peer evaluator
CLASSROOM OBSERVATIONS	
<i>Who are the observers who rate teachers' effectiveness?</i>	<p>HCPS's teacher evaluation system uses three types of observers:</p> <ul style="list-style-type: none"> • Principals <ul style="list-style-type: none"> - District subject-area supervisors conduct some observations for struggling teachers - Assistant principals assist with classroom observations (the principal is required to do a minimum of one formal observation for each teacher each year) • Peer evaluators (for experienced teachers) <ul style="list-style-type: none"> - Applicants chosen by a committee that includes directors, principals, supervisors, teachers, and union representatives, based on evaluation results - Peer evaluators must have strong evaluation results and at least five years of teaching experience - Peer evaluators are from the teacher's subject area whenever possible - Being a peer evaluator is a full-time position (these teachers do not continue teaching while serving in this capacity) • Mentor evaluators (for new teachers) <ul style="list-style-type: none"> - Applicants chosen by a committee that includes directors, principals, supervisors, teachers, and union representatives, based on communication skills, mentoring experience, leadership, and knowledge of pedagogy and instructional practices - Mentors must have at least five years of teaching experience and strong evaluation results - Each mentor is expected to work with 15 new teachers - Mentors will not be content-specific, but will be specific by grade level (elementary, middle, high school) • HCPS received more than 660 applicants for 116 mentor/peer teacher positions in its first year (2010–11) and more than 500 applicants for an additional 70 positions in year two (2011–12)
<i>What training do observers receive?</i>	<ul style="list-style-type: none"> • The Cambridge Education Group and New Teacher Center provided initial training. During the 2011–12 school year, the district's professional development department assumed responsibility for primary training • Training based on Charlotte Danielson's Framework for Teaching • Principal and peer/mentor evaluators participate in 4 to 10 days of training • Early in their training, evaluators also conduct paired observations with another evaluator • Mentors are coached on a regular basis by lead mentors • Each observer is calibrated by Cambridge Education on an annual basis
<i>Are observers compensated for this work?</i>	Peer/mentor evaluator positions are full-time jobs that are compensated accordingly. Teachers earn their regular salary plus a \$5,000 stipend.

SITE PROFILES

<p><i>How often are teachers observed?</i></p>	<p>Each teacher is observed a minimum of five times a year, with struggling teachers receiving more observations. For example:</p> <ul style="list-style-type: none"> • An experienced teacher with a high evaluation score will have two formal and three informal observations • At the other end of the continuum, a teacher designated as unsatisfactory will have seven formal and four informal observations
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<p>Observations are ongoing from the third week of school until mid to late May.</p>
<p><i>Are observations announced or unannounced?</i></p>	<ul style="list-style-type: none"> • In year one (2010–11) all observations were announced • In year two (2011–12), informal observations are unannounced
<p><i>What is the basis for the observation rubric?</i></p>	<p>HCPS uses an observation rubric that is a modified version of <u>Charlotte Danielson's Framework for Teaching</u>.</p>
<p><i>What is included in the observation scoring rubric?</i></p>	<p>HCPS's observation rubric has ratings for:</p> <ul style="list-style-type: none"> • Planning and preparation • Classroom environment • Instruction • Professional responsibilities (assessed by the principal only)
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • Mentor/peer evaluators are also evaluated. Their rating includes scores from: <ul style="list-style-type: none"> - Program director who oversees mentor and peer evaluators (60%) - Teachers with whom they work and observe (40%) • Five percent of the evaluation score that principals receive is based on their evaluation of teachers, which includes how strongly their ratings are correlated with those of peers and mentors, as well as teachers' value-added scores
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • Yes, a conference follows every formal observation • Observers also provide written feedback after each informal observation

ACCURACY, VALIDITY, AND RELIABILITY	
<i>Is there a process to validate non-standardized assessments?</i>	Not applicable (all assessments are standardized)
<i>Are there procedures to ensure that observation scores are valid?</i>	<ul style="list-style-type: none"> • The district has hired Cambridge Education to calibrate observers • Peers and mentors frequently observe together as a calibration exercise; principals will sometimes do the same • Observation data are reviewed at the district level on a weekly basis, and “outliers” are identified for additional calibration
<i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i>	HCPS conducted an end-of-year analysis and found the correlation between written and value-added scores to be approximately 0.36, representing fairly strong and positive alignment.
<i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i>	Not applicable. HCPS’s evaluation system includes two measures: student achievement and teacher observation.
USING AND REPORTING EVALUATION RESULTS	
<i>What are the rating categories?</i>	Teachers will receive an overall rating of 1 through 5 (to be introduced in 2011–12).
<i>How are results reported?</i>	<ul style="list-style-type: none"> • Individual teacher results are reported to teachers and administrators only • Starting in 2013–14, HCPS will implement a new salary system in which a teacher’s salary level will reflect evaluation results from the past two years, with teachers receiving a pay increase if they achieve a higher rating for two consecutive years • Teachers employed by the 2009–10 school year may choose to remain on the current salary schedule
<i>Are there consequences tied to evaluation results? If yes, what are they?</i>	<p>Yes, evaluation results are tied to professional development, teacher pay, increased responsibility, and dismissal.</p> <ul style="list-style-type: none"> • Professional development/improvement plan <ul style="list-style-type: none"> - Teachers who need to improve in a particular focus area (as demonstrated through their evaluation) will receive targeted professional development - The number of observations a veteran teacher receives will reflect his/her previous evaluation score(s) • Pay <ul style="list-style-type: none"> - Starting in 2013–14, a teacher’s career ladder level, and corresponding pay level, will reflect evaluation data from the past two years. Career ladder levels are not necessarily associated with additional responsibility (although several positions, including peer and mentor evaluator positions, are), but will be rewarded with additional pay - Teachers may move up the career ladder and receive higher pay if they receive a higher performance evaluation for two consecutive years - Similarly, a teacher may move down the career ladder and receive decreased pay if her annual evaluation level decreases for three years in a row

SITE PROFILES

<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<ul style="list-style-type: none"> • Increased responsibility <ul style="list-style-type: none"> - HCPS will implement a performance-based career ladder in 2013–14 that creates a new set of steps for teacher leaders and mentor/peer evaluators, which include increased responsibilities and pay (see slide 6) - Teachers will eventually have to score a 4 or 5 to qualify for these new positions. • Dismissal <ul style="list-style-type: none"> - Teachers who receive a 1 or 2 rating for two consecutive years or more and do not take advantage of feedback and professional development resources could be recommended for dismissal by the principal - The dismissal process has not changed: probationary teachers (teachers in their first three years) may not be renominated at the end of the year, but may request a hearing similar to that those given to tenured teachers
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<p>Yes, but teachers can appeal their ratings based only on considerations of procedure and data (whether the correct students/scores were used). An evaluator’s judgment cannot be the basis of an appeal.</p>
<p><i>If performance results are available, what is the score distribution?</i></p>	<ul style="list-style-type: none"> • HCPS’s 1–5 rating system is being implemented in 2011–12 • In 2010–11, scores were on a scale of 0–100, with a median score of 61.5
<p>TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT</p>	
<p><i>What role did teachers play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • Teachers and leaders of the Hillsborough Classroom Teachers Association constitute approximately half of the district’s Teacher Evaluation Committee, which designed EET and meets each year to discuss and revise the system • Staff from HCPS always present major decisions and policies to the Teacher Advisory Committee, which meets monthly to provide feedback on EET
<p><i>What role did the teachers union play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • Both the local union president and president of the national American Federation of Teachers backed Hillsborough’s plan • The union president is a member of the district’s steering committee on evaluation
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<ul style="list-style-type: none"> • District-wide teacher and principal evaluation committees meet at the end of each year to discuss what worked as well as identify areas for improvement • Hillsborough County has committed to revising its teacher evaluation system “as often as necessary”

GLOSSARY

<p><i>Mentor evaluator</i></p>	<ul style="list-style-type: none"> • Full-time mentors for new (first- and second-year) teachers who also evaluate new teachers (but not the teachers they mentor) • Mentor teachers must demonstrate strong communication skills, mentoring experience, leadership, and knowledge of pedagogy and instructional practices • Mentors must have at least five years of teaching experience and a history of high performance in the classroom • Each mentor is expected to work with 20 first- and second-year teachers • Not subject-specific, but grade-level specific when possible
<p><i>Peer evaluator</i></p>	<ul style="list-style-type: none"> • Teachers chosen to be full-time peer evaluators based on their own previous evaluations, interpersonal skills, etc. • Must have at least five years of teaching experience • Content-specific whenever possible

TIMELINE

Not available.

SOURCES

Curtis, Rachel. (2012, March). *Building it together: The design and implementation of Hillsborough County Public Schools’ teacher evaluation system*. Retrieved from <http://www.aspeninstitute.org/publications/hillsborough-county-public-schools>

Hillsborough County Public Schools. *Empowering effective teachers*. Retrieved from <http://communication.sdhc.k12.fl.us/empoweringteachers/>

Hillsborough County Public Schools. *Multiple measures of teacher effectiveness in Hillsborough County Public Schools: Implementing value-added measures*. Retrieved from <http://communication.sdhc.k12.fl.us/eethome/casestudies/Implementing%20Value-added%20Measures%20in%20Hillsborough%20County.pdf>

HOUSTON, TX

TEACHER APPRAISAL AND DEVELOPMENT SYSTEM

OVERVIEW

In 2010, Houston Independent School District (HISD) engaged in a district-wide effort to design a plan, called the Strategic Direction, for transforming the district. HISD’s Strategic Direction involves five core initiatives, one of which is placing an effective teacher in every classroom. This initiative, in part, included the design of a new teacher appraisal and development system, which HISD implemented for the first time in the 2011–12 school year.

HISD’s appraisal system currently contains three main components: instructional practice, professional expectations, and student performance. The weight of each component varies based on scores in the other areas, but student performance accounts for roughly half of a teacher’s evaluation score. Appraisers evaluate teachers on instructional practice and professional expectations using a standard rubric of performance criteria administered through multiple, unannounced classroom observations conducted by school leaders during the year. Appraisers also select the measures for the student performance component from a district-approved list based on the teacher’s subject and grade. All teachers must have a minimum of two student performance measures for student performance to be included in their appraisal rating.

In 2011–12, the appraisal system offered five types of student performance measures, including value-added growth, comparative growth on district assessments, and progress on end-of-course (EOC) assessments or student work, and, in a very limited number of instances, attainment on EOC assessments. In order to reduce the variability in measures from teacher to teacher, HISD is developing more than 20 end-of-course assessments that will measure growth and attainment in currently untested subjects.

Texas has used its current statewide teacher evaluation system since the 1997–98 school year. Houston, however, wanted to use a different model. To do so, state policy required that HISD involve stakeholders at every level of the system’s development. As a result, each component was drafted based on recommendations from school-based committees, opened for public comment, and reviewed by the district-level advisory committee. Although challenging, frequent and transparent communication with stakeholders has been a priority for Houston throughout the development of the system.

Houston was in the process of refining its teacher evaluation system when this document was published. The web version of this publication will link to the most recent supporting documents as they become available throughout the spring and summer.

ABOUT THE SCHOOL SYSTEM

<i>Number of students</i>	202,773
<i>Number of schools</i>	298
<i>Number of teachers</i>	12,829

BACKGROUND

<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	No

<i>In effect since</i>	2011–12 (excluding student performance component)
<i>Are there plans for additional phases/components?</i>	Although Houston collects data for the Student Performance component, those scores will not be included in a teacher’s summative rating until 2012–13.
<i>Who gets evaluated?</i>	All teachers
<i>How often do teachers receive a rating?</i>	<ul style="list-style-type: none"> • Midyear rating (formative) • End-of-year rating (summative of Instructional Practice and Professional Expectations) • Final rating at beginning of next school year (which includes student performance) • Only final rating counts for consequences
<i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i>	Not applicable
STUDENT ACHIEVEMENT MEASURES	
<i>Assessments to measure student achievement</i>	<ul style="list-style-type: none"> • <u>State of Texas Assessments of Academic Readiness (STAAR)</u> <ul style="list-style-type: none"> - Reading: grades 3–8 - Math: grades 3–8 - Science: grades 5, 8 - Social studies: grade 8 - English I, II, III - Algebra I, II, geometry - Biology, chemistry, physics - World geography, world history, U.S. history, world geography - STAAR Modified or STAAR Alternate Assessment: special education students • Texas English Language Proficiency Assessment System (TELPAS): Grades K–12 • <u>Stanford 10</u> and <u>APRENDA 3</u> Achievement Tests (norm-referenced) <ul style="list-style-type: none"> - English language arts/Secondary language acquisition: grades 1–8 - Reading: grades 1–8 - Math: grades 1–8 - Science: grades 3–8 - Social studies: grades 3–8 • AP/IB exams: high school • CTE courses with certification exams: high school
<i>Are any new assessments in development to use with teacher evaluations?</i>	Yes, HISD is creating end-of-course common assessments for priority untested courses that measure growth and attainment.
<i>What academic/achievement criteria are included?</i>	Beginning in 2012–13, every teacher will have at least two Student Performance measures out of five possible types of measures: value-added growth; comparative growth; fall-to-spring student progress on approved assessments; fall-to-spring student progress on performance tasks or work products; and student attainment on approved assessments.

SITE PROFILES

<p><i>What academic/achievement criteria are included?</i></p>	<ul style="list-style-type: none"> • Courses that each teacher teaches will determine required measures. Otherwise, appraisers will determine which courses will be measured for student performance and which student performance measures will be used for each course (see list of anticipated measures, pp. 8–16) • Value-added growth <ul style="list-style-type: none"> - In grades/subjects or courses where available, one measure must be value-added growth using the <u>Educational Value-Added Assessment System (EVAAS)</u> - EVAAS uses up to four years of test scores to compare a student’s actual growth to his projected growth based on the average growth of similar students across the state - Every teacher receives a Teacher Gain Index (TGI) score that shows how much her students grew academically in relation to the expected district average growth for that grade and subject, accounting for measurement error - The TGI score is then translated into a performance score between 1 and 5, 1 meaning that the teacher’s students on average fell substantially short of expected average growth, and 5 meaning that the teacher’s students on average substantially exceeded expected average growth; a score of 3 means that students made expected average growth - Currently about 25 percent of HISD teachers have value-added data • Comparative growth on district-wide EOC/EOY assessments <ul style="list-style-type: none"> - HISD has developed a comparative growth model for use in the appraisal system - Growth will be measured by comparing results among similar groups of students - Students who ended the previous year with a similar result are compared to one another the following year - Teachers will get a rating that measures the extent to which their students meet or exceed the growth of similar students • Student progress on district-wide OR appraiser-approved EOC/EOY assessments <ul style="list-style-type: none"> - Appraiser will select district-wide assessments OR teacher will select or create assessments, and appraiser will approve them - Teacher will set ambitious and feasible targets for students based on starting points, and appraiser will approve them - Appraiser will give teacher rating based on percentage of students who reach targets by the end of the year (<u>see rubric, p. 7</u>) • Student progress using EOC/EOY performance tasks and work products <ul style="list-style-type: none"> - Teachers will select student performance tasks and work products, and appraiser will approve them <ul style="list-style-type: none"> » Example of performance task: a recital in music class » Example of work product: a portfolio in art class - Teacher will set ambitious and feasible targets for students based on starting points, and appraiser will approve them - Appraiser will give teacher rating based on percentage of students who reach targets by the end of the year (<u>see rubric, p. 7</u>) • For Pre-K only: Student attainment on appraiser-approved or district-wide EOC/EOY assessment <ul style="list-style-type: none"> - Appraiser will select district-wide Pre-K assessments - Appraiser will give teacher rating based on percentage of students meeting district standard on assessments <p>For more information on Student Performance measures, see here.</p>
--	--

<p><i>How much do the student achievement measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • Teachers receive a score of 1 to 4 in each of the components • Final ratings are calculated through a matrix. Therefore, the weight of each component varies depending on the teacher's performance on other pieces • Student performance accounts for roughly half of a teacher's final evaluation score
<h2>NONACADEMIC MEASURES</h2>	
<p><i>What nonacademic criteria are included?</i></p>	<p>The two nonacademic components are Instructional Practice and Professional Expectations.</p> <ul style="list-style-type: none"> • <u>Instructional Practice (IP)</u> <ul style="list-style-type: none"> - Criteria include using data to inform instruction; designing effective lesson plans and assessments; differentiating for student needs; checking for understanding; promoting high academic expectations; and implementing effective classroom management • <u>Professional Expectations (PE)</u> <ul style="list-style-type: none"> - Criteria include respecting colleagues; complying with teacher attendance policies; dressing professionally; communicating with parents; and participating in professional development
<p><i>How much do the nonacademic measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • Teachers receive a score of 1 to 4 in each of the components • Final ratings are calculated through a matrix. Therefore, the weight of each component varies depending on the teacher's performance on other pieces • Instructional Practice and Professional Expectations combined account for roughly half of a teacher's final evaluation score • Instructional Practice weighs more heavily than Professional Expectations
<h2>CLASSROOM OBSERVATIONS</h2>	
<p><i>Who are the observers who rate teachers' effectiveness?</i></p>	<ul style="list-style-type: none"> • Assistant principals and deans make up the majority • Out of 800 current appraisers, only 300 are principals
<p><i>What training do observers receive?</i></p>	<p>All appraisers receive training from HISD, which includes:</p> <ul style="list-style-type: none"> • Spring training on student performance, which includes eight online training sessions that total four hours, followed by two online assessments that appraisers must pass to receive certification • Two-day summer training on the appraisal and development system, followed by two assessments on student performance and instructional practice that appraisers must pass to receive certification • HISD provides ongoing training and support for appraisers during the summer and throughout the school year, including skill-building sessions, guides, and videos that provide guidance on leading conferences throughout the year
<p><i>Are observers compensated for this work?</i></p>	<p>No, observations are part of the administrator's job description/salary.</p>

SITE PROFILES

<i>How often are teachers observed?</i>	<ul style="list-style-type: none"> • Longer classroom observations (at least 30 minutes each)—at least two annually • Shorter walkthroughs (at least 10 minutes each)—at least two annually • Appraisers must conduct enough observations to observe all of the rubric criteria
<i>Is there a time frame within which observations must be conducted?</i>	No, but HISD encourages appraisers to conduct at least one longer observation and one shorter walk-through prior to midyear progress conference.
<i>Are observations announced or unannounced?</i>	Unannounced
<i>What is the basis for the observation rubric?</i>	<ul style="list-style-type: none"> • HISD formed a working group of HISD teachers, administrators, and staff to develop its observation rubric • The group reviewed many existing rubrics, including Danielson and TAP, but ultimately it is a homegrown rubric that draws on strengths of various existing frameworks
<i>What is included in the observation scoring rubric?</i>	Appraisers observe teachers using the Instructional Practice rubric and Professional Expectations rubric .
<i>How are observers held accountable for their scoring?</i>	<ul style="list-style-type: none"> • HISD monitors teacher ratings and data gathered from conferences throughout the year, and compiles information for district and school leaders • Staff Review Process: Principals discuss teacher ratings and data with their managers and with their staff
<i>Do teachers have an opportunity to debrief with observers?</i>	<ul style="list-style-type: none"> • Longer observations require in-person and written feedback within 10 days • Shorter walk-throughs require written feedback within 10 days • In addition, appraisers and teachers meet three times a year to set goals, review student data, and discuss next steps for improvement
ACCURACY, VALIDITY, AND RELIABILITY	
<i>Is there a process to validate non-standardized assessments?</i>	<ul style="list-style-type: none"> • All student performance measures must be cumulative end-of-year assessments • Appraisers validate EOC/EOY Assessments, performance tasks, and work products using district-wide criteria (see p. 18) that evaluate the scope, rigor, format, and reliability of an assessment
<i>Are there procedures to ensure that observation scores are valid?</i>	No, but raters are held accountable for rating quality—see <i>“How are observers held accountable for their scoring?”</i>
<i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i>	HISD is tracking data in the first year of implementation. More formal reviews will be in place once initial data are available.
<i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i>	See above

USING AND REPORTING EVALUATION RESULTS	
<i>What are the rating categories?</i>	Teachers will receive an overall rating of ineffective, needs improvement, effective, or highly effective.
<i>How are results reported?</i>	Individual teacher results are reported to teachers and administrators only. <ul style="list-style-type: none"> • Teachers are notified of ratings at the end of the school year • HISD is developing an online data portal where teachers and administrators can access evaluation data and results throughout the year
<i>Does the system allow teacher evaluation results to be reported by the training programs teachers attended?</i>	<ul style="list-style-type: none"> • Yes, Houston's system has the ability to report results by teacher training program, although HISD has not decided if it will do so publicly • As part of its recruitment strategy, HISD plans on using 2011–12 data from the evaluation system to determine which teacher preparation programs to recruit more heavily from
<i>Are there consequences tied to evaluation results? If yes, what are they?</i>	<p>Yes, evaluation results are tied to professional development, dismissal, and reduction-in-force decisions.</p> <ul style="list-style-type: none"> • Professional development/improvement plan <ul style="list-style-type: none"> - Teachers who are not meeting Instructional Practice expectations may be placed on a specific improvement plan, called a Prescriptive Plan for Assistance, at the appraiser's discretion - Teachers are placed on and off specific improvement plans at appraiser's discretion - Improvement plans include specific goals set by the teacher and the appraiser for instructional improvement by a particular date • Dismissal <ul style="list-style-type: none"> - Evaluation results are not directly tied to dismissal, but teachers may be dismissed based on consistently poor student growth data - Teachers may also be recommended for dismissal based partly on evaluation results • Other <ul style="list-style-type: none"> - Evaluation results are used in reduction-in-force decisions <p>Note: HISD is in the design phase of a new teacher compensation and career pathways system, which will use evaluation results to determine pay and opportunities for increased responsibility and to reach more students.</p>
<i>Can teachers appeal their ratings? If yes, how?</i>	<ul style="list-style-type: none"> • Any teacher may appeal her ratings only once: either the midyear formative rating or end-of-year rating • Teachers may request a second appraiser if they disagree with their final Instructional Practice, Professional Expectations, and/or Student Performance ratings • The second appraiser will be appointed by the superintendent or his/her designee. Teachers may request that the second appraiser have content knowledge specific to the area being taught • The second appraiser will conduct an unannounced classroom observation and review all materials related to the disputed areas, then assign a second rating to the teacher based upon that information • The second rating will then be averaged with the original appraiser's rating
<i>If performance results are available, what is the score distribution?</i>	Data are unavailable.

SITE PROFILES

TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT	
<i>What role did teachers play in developing the evaluation system?</i>	<ul style="list-style-type: none"> • In 2010–11, a seven-month development period involved more than 2,600 teachers. Process involved: <ul style="list-style-type: none"> - More than 250 School-based <u>Shared Decision-Making Committees</u> (SDMCs) made up of teachers, school administrators, parents, and community members - <u>District Advisory Committee</u> (DAC), made up of school-elected teachers, principals, district administrators, parents, and community members - Working groups made up of teachers that met regularly to build tools and rubrics to be used in implementation • SDMCs submitted recommendations on appraisal criteria and process to the DAC • Two-week public comment period gave additional stakeholders opportunity to make recommendations • DAC oversaw the final product to be submitted for approval by the board of education
<i>What role did the union play in developing the evaluation system?</i>	Representatives from the teachers union participate on SDMCs, DAC, and working groups.
<i>Is there a formal process for revising the evaluation system over time?</i>	<ul style="list-style-type: none"> • State education code requires annual formal evaluation of system to make changes over time • Changes to evaluation process and criteria will be reviewed by school-based <u>Shared Decision-Making Committees</u> (SDMCs) and <u>District Advisory Committee</u> (DAC), then presented to the board of education for approval
GLOSSARY	
<i>Appraiser</i>	<ul style="list-style-type: none"> • Principals, assistant principals, and deans who are responsible for helping teachers select growth measures, evaluating teachers, and rating them based on results • All appraisers receive training from the district
<i>Comparative growth</i>	<ul style="list-style-type: none"> • How much a teacher’s students grew academically in comparison to similar students
<i><u>District Advisory Committee (DAC)</u></i>	<ul style="list-style-type: none"> • District-level committee made up of school-elected teachers, principals, district administrators, parents, and community members • During development of teacher appraisal system, oversaw the final product to be submitted for approval by the board of education • All departures from criteria and process of the statewide evaluation system will be reviewed by the DAC before being presented to the board of education for approval

<p><u>School-based Shared Decision-Making Committee (SDMC)</u></p>	<ul style="list-style-type: none"> • School-level committee made up of teachers, principals/other school leaders, parents, and community members • During development of teacher appraisal system, submitted recommendations on appraisal criteria and process to the District Advisory Committee (DAC) • State education code requires involvement of SDMC in all departures from the criteria and process of the statewide evaluation system
<p><i>Value-added growth</i></p>	<ul style="list-style-type: none"> • How much a teacher’s students grew academically in relation to the expected average growth for that particular group of students

TIMELINE



SITE PROFILES

SOURCES

Houston Independent School District. (2011). *Appraisal and Development System Proposal Materials*. Retrieved May 3, 2012, from <http://hisdeffectiveteachers.org/final-appraisal-and-development-system-proposal>

Houston Independent School District. (2011). *Building a better teacher appraisal and development system: Summary*. Retrieved November 25, 2011, from http://hisdeffectiveteachers.org/assets/Appendix_1_Summary_of_the_Appraisal_and_Development_System.pdf

Houston Independent School District. *Facts and figures about HISD*. Retrieved November 25, 2011, from <http://www.houstonisd.org/HISDConnectDS/v/index.jsp?vgnextoid=62c6757761efc010VgnVCM10000052147fa6RCRD&vgnextchannel=2e2b2f796138c010VgnVCM10000052147fa6RCRD>

Houston Independent School District. *HISD teacher appraisal and development system instructional practice and professional expectations rubrics*. Retrieved November 25, 2011, from <http://www.houstonisd.org/HumanResources/Home/Appraisals%20Systems/Teacher/2011-2012%20HISD%20Teacher%20IP%20and%20PE%20Rubrics.pdf>

Houston Independent School District. *HISD's teacher appraisal and development system: System overview*. Retrieved November 25, 2011, from <http://www.houstonisd.org/HumanResources/Home/Appraisals%20Systems/Teacher/110902%20General%20Appraisal%20and%20Development%20Overview.pdf>

Texas Association of School Boards. (2011). *The ratings game: Can we fix Texas' broken teacher evaluation system?* Retrieved May 1, 2012, from http://www.tasb.org/services/hr_services/hrexchange/2011/September11/teacher_eval_lead.aspx

NEW HAVEN, CT

TEACHER EVALUATION AND DEVELOPMENT PROCESS (TEVAL)

OVERVIEW

In 2009, New Haven Public Schools (NHPS) launched a district-wide reform effort called New Haven School Change Initiative, which included revamping how it manages its teaching talent. That fall, NHPS and the New Haven Federation of Teachers submitted recommendations to the board of education to reform the teacher evaluation system. The result is the Teacher Evaluation and Development Process, or TEVAL.

Under the new system, at least half of a teacher's rating must reflect student achievement, measured as progress toward "growth goals." Teachers work with administrators to identify at least two student growth goals that reflect ambitious but reasonable growth for students in her class. The other half of a teacher's final rating is based on nonacademic measures, including instructional practice and professional values. Although administrators are primarily responsible for rating teachers on these nonacademic measures, third-party evaluators employed by the district conduct an additional observation for teachers rated in the lowest and highest categories ("needs improvement" and "exemplary") as a way to validate the results.

NHPS is working to ensure reliability of the assessments used to measure student growth, particularly for untested subjects and grades. The district also continues to work with teachers to set appropriate student growth goals.

ABOUT THE SCHOOL SYSTEM

<i>Number of schools</i>	45
<i>Number of students</i>	20,759
<i>Number of teachers</i>	1,850

BACKGROUND

<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	No
<i>In effect since</i>	2010–2011
<i>Are there plans for additional phases / components?</i>	No
<i>Who gets evaluated?</i>	All teachers, school psychologists, social workers, principals
<i>How often do teachers receive a rating?</i>	Once a year
<i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i>	Not applicable

SITE PROFILES

STUDENT ACHIEVEMENT MEASURES	
<i>Assessments to measure student achievement</i>	<p>State tests include:</p> <ul style="list-style-type: none"> • Connecticut Mastery Test (CMT) <ul style="list-style-type: none"> - Grades: 3 through 8 - Subjects: reading, mathematics, writing for all grades; science for grades 5-8 • Connecticut Academic Performance Test (CAPT) <ul style="list-style-type: none"> - Grade 10 - Subjects: reading, mathematics, writing, science • CMT/CAPT Modified Assessment System (MAS) and CMT/CAPT Skills Checklist <ul style="list-style-type: none"> - Grades 3-8, 10 Special Education - Subjects: reading, mathematics, writing, science • <u>Developmental Reading Assessment</u>
<i>Are any new assessments in development to use with teacher evaluations?</i>	<p>The state of Connecticut is currently participating in the <u>SMARTER Balanced Assessment System</u> aligned to Common Core Standards to be implemented 2014-2015</p> <ul style="list-style-type: none"> • Grades: 3-8, 11 • Subjects: Mathematics and ELA • Will be used to measure growth and achievement
<i>What academic / achievement criteria are included?</i>	<p>Teachers are assessed on progress toward <u>growth goals</u> based on CMT data or other approved assessments.</p> <ul style="list-style-type: none"> • Teachers work with their instructional managers (IM) to set a minimum of two and a maximum of four <u>growth goals</u> for students each year. The IM is the principal, assistant principal, or other administrative leader accountable for a teacher's evaluation and development <ul style="list-style-type: none"> - For CMT-tested grades and subjects, one growth goal must be based on CMT data - Other goal(s) may be based on IM-approved assessments such as: <ul style="list-style-type: none"> » District-wide assessments, including: <ul style="list-style-type: none"> › Grades 7–8, and high school: Quarterly assessments by course in English, math, science, social studies, and world languages › Grades 3–8: <ul style="list-style-type: none"> – <u>DRA2+</u> or <u>DRP</u> – Pre- and post- District Connecticut Mastery Test (DCMT) in language arts and math – Mini-assessments in math and language arts › Grades K–2: <u>DRA2+</u>, phonemic awareness, oral language, and math » Student work portfolios that assess district-wide "<u>21st Century Competencies</u>," six competencies designed by a committee of teachers and district officials that assess a student's critical thinking, creativity, initiative, citizenship, and technological research skills » Teacher-created assessments (Note: scoring on teacher-created assessments and student work will be validated, whenever possible, through teacher scoring groups, spot-checking, or some other mechanism) - <i>Example of elementary math goal: The average vertical scale score of the class on CMT math will increase by 25 points</i> - <i>Example of high school art goal: 90 percent of final portfolio presentations attain the level of "shows significant improvement" or "shows outstanding improvement" on rubric at year-end review</i> • Teachers use previous year's data or their own pre-assessment of student performance as the baseline for setting growth goals

<p><i>How much do the student achievement measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • Teachers receive a Student Learning Growth score between 1 and 5 • Combined Instructional Practice and Professional Values score is then combined with Student Learning Growth score to calculate an overall rating using a matrix (See p. 7 of conference form) <ul style="list-style-type: none"> - A teacher cannot be rated exemplary overall unless she is rated strong or exemplary on student learning growth
<p>NON-ACADEMIC MEASURES</p>	
<p><i>What non-academic criteria are included?</i></p>	<p>TEVAL has two nonacademic components: Domains of instructional practice and professional values.</p> <ul style="list-style-type: none"> • Instructional Practice Performance Continuum criteria are: <ul style="list-style-type: none"> - Planning and Preparation <ul style="list-style-type: none"> » Establishes student learning goals and objectives; designs units, lessons, assessments; addresses social and emotional needs of students - Classroom practice <ul style="list-style-type: none"> » Differentiates instruction; monitors understanding; engages students; incorporates inquiry and critical thinking - Reflection and use of data <ul style="list-style-type: none"> » Uses assessments to inform instruction; reflects on effectiveness and student interaction to make needed adjustments • Professional Values Framework and Performance Continuum criteria are collaboration and collegiality; self-improvement; reliability; high expectations; respect; responsiveness and outreach; professionalism and judgment
<p><i>How much do the non-academic measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • Teachers are assigned a combined Instructional Practice and Professional Values score between 1-5 <ul style="list-style-type: none"> - Instructional Practice accounts for 80% of the combined score - Professional Values accounts for 20% of the combined score • Combined Instructional Practice and Professional Values score is then combined with Student Learning Growth score to calculate an overall rating using a matrix (See p.7 of conference form) <ul style="list-style-type: none"> - A teacher cannot be rated overall Exemplary unless she is rated Strong or Exemplary on student learning growth

SITE PROFILES

CLASSROOM OBSERVATIONS	
<i>Who are the observers who rate teachers' effectiveness?</i>	<p>There are two kinds of observers in TEVAL: Instructional Managers and Third Party Validators.</p> <ul style="list-style-type: none"> • Each teacher is observed and rated by his/her Instructional Manager (IM); IMs are often principals, assistant principals, or other administrative leaders • To help ensure validity and fairness, teachers who are identified at the highest rating (Exemplary) or lowest rating (Needs Improvement), are also observed and rated by an IM by a Third Party Validator (TPV), former teachers who are not currently affiliated with the district, but have demonstrated effectiveness in the classroom in the past
<i>What training do observers receive?</i>	<ul style="list-style-type: none"> • IMs and TPVs receive training from the NHPS central office • Training includes practicing goal-setting, observations, gathering evidence, and rating teacher practice through videos • IMs receive 2 days of training in the summer on instructional practice • TPVs receive a 4-hour training at the beginning of the year, and at least one session mid-year
<i>Are observers compensated for this work?</i>	<p>Yes, TPVs receive \$225 per observation and written report. IMs do not receive any additional compensation as conducting evaluations is part of their job responsibilities.</p>
<i>How often are teachers observed?</i>	<ul style="list-style-type: none"> • For all teachers: <ul style="list-style-type: none"> - IMs conduct informal observations about every two weeks (not limited to classroom) - At least one a year, but IMs may conduct formal observations at their discretion • For Exemplary teachers: In addition to IM observations, TPVs conduct two additional formal observations • For Needs Improvement teachers: In addition to IM observations, TPVs conduct three additional formal observations
<i>Are observations announced or unannounced?</i>	<ul style="list-style-type: none"> • For all teachers: IM formal observations are announced • For Exemplary teachers: Additional TPV observations will include one announced, one unannounced • For Needs Improvement teachers: Additional TPV observations will include one announced and two unannounced
<i>What is the basis for the observation rubric?</i>	<p>NHPS reviewed several rubrics when designing the rubric, including Charlotte Danielson's <i>Framework for Teaching</i>.</p>

<p><i>What is included in the observation scoring rubric?</i></p>	<p>Evaluators use the Instructional Practice Performance Continuum and Professional Values Framework and Performance Continuum when observing teachers. See “<i>Non-academic measures</i>” for details.</p>
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • TPVs conduct additional observations for Exemplary and Needs Improvement teachers to ensure validity of those scores from IMs • When teachers and IMs disagree on observation results or topics discussed during conferences, the teachers union will set up meetings with the district, teacher and IM to discuss
<p><i>Do teachers have an opportunity to debrief the observations with observers?</i></p>	<ul style="list-style-type: none"> • IMs frequently debrief informal observations with teachers in the most convenient way (verbal, written, e-mail, etc.) • TPVs must provide written feedback for teachers within 48 hours of observation; TPVs must then complete a review using rubrics to rate teachers within 72 hours of observation • Additionally, all IMs and teachers have at least three conferences annually for goal-setting and progress assessment
<p>ACCURACY, VALIDITY, AND RELIABILITY</p>	
<p><i>Is there a process to validate non-standardized assessments?</i></p>	<p>There is not a formal process to validate non-standardized measures, but the district recommends several checks, including:</p> <ul style="list-style-type: none"> • Having literacy coaches spot check teacher administered assessments • Having teachers administer the assessments for different classes (not their own)
<p><i>Are there procedures to ensure that observation scores are valid?</i></p>	<ul style="list-style-type: none"> • For NI and Exemplary teachers, a TPV is used • Thus far, there has been strong alignment between IM and TPV ratings
<p><i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i></p>	<ul style="list-style-type: none"> • If a teacher’s score on “instructional practice and professional values” is too far out of line with the teacher’s score on “student learning growth,” the rating is subject to review by NHPS central office • Central office will work with IM to look at growth goals and observation records to find roots of discrepancies <ul style="list-style-type: none"> - To date, the most common issue has been that the goal set was not rigorous enough or it was too ambitious • Associate superintendent will make final decision on teacher rating when there are concerns
<p><i>Is there a procedure to validate that other non-academic measures are correlated with student outcomes?</i></p>	<p>Yes, see above.</p>

SITE PROFILES

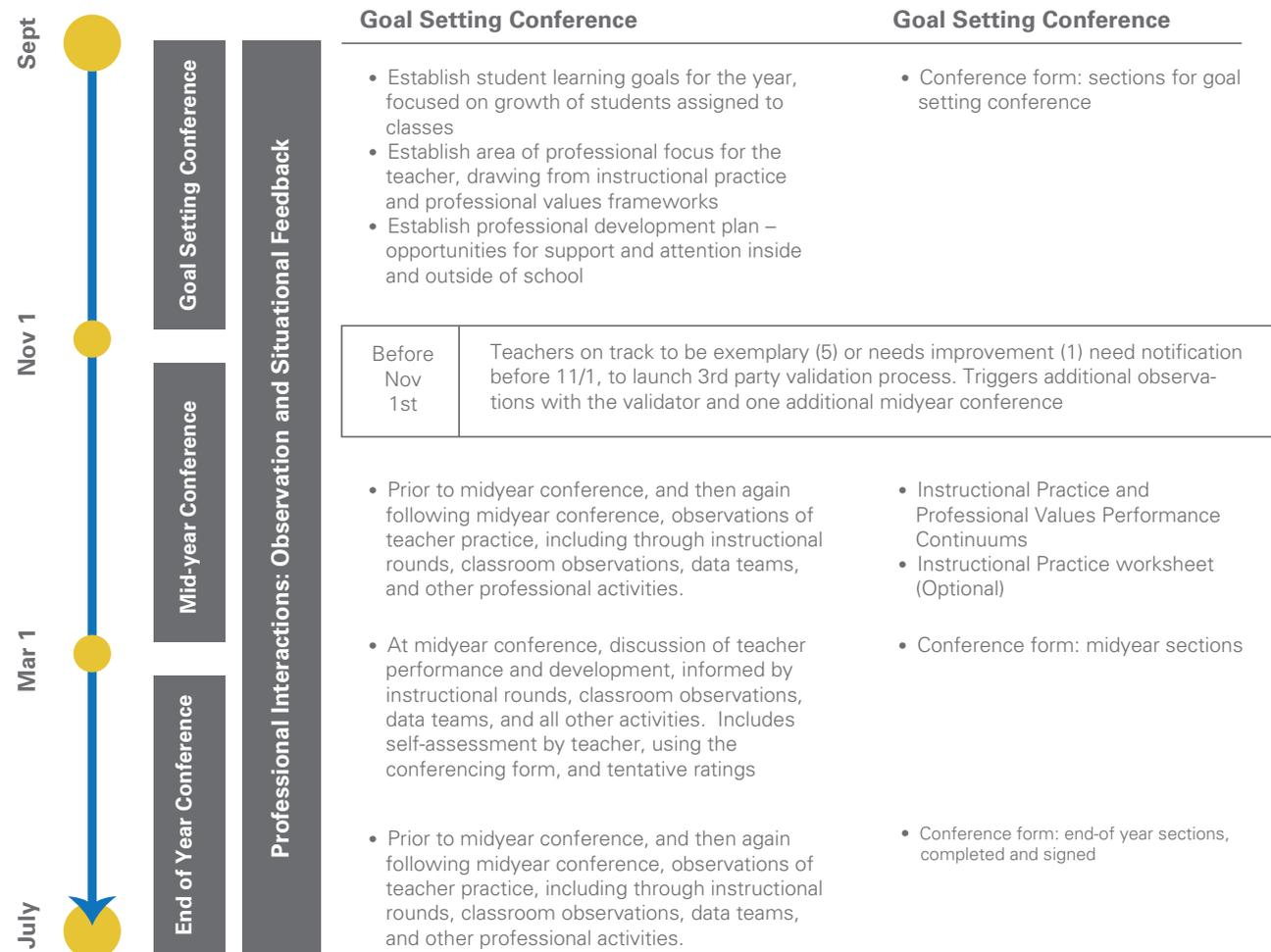
USING AND REPORTING EVALUATION RESULTS	
<i>What are the rating categories?</i>	<p>Teachers will receive an overall rating of 1-5:</p> <ul style="list-style-type: none"> • Exemplary (5) • Strong (4) • Effective (3) • Developing (2) • Needs Improvement (1)
<i>How are results reported?</i>	<ul style="list-style-type: none"> • Individual teacher results are reported to teachers and administrators only • A recent report titled "Talent Management Update" is publicly available on the district's website, and shows district-wide results
<i>Are there consequences tied to evaluation results? If yes, what are they?</i>	<p>Yes, evaluation results are tied to increased responsibility, tenure, professional development, and dismissal.</p> <ul style="list-style-type: none"> • Professional development/improvement plan <ul style="list-style-type: none"> - "Strong" and "effective" teachers develop a professional development plan to improve skills and become exemplary. They also receive targeted PD administered by the school - Teachers on track to receive a needs improvement rating must be identified and notified by November 1 to ensure that teachers have enough time to improve, and that potential dismissals may occur in the same year they are designated (IMs may use previous year's ratings to make early, targeted observations of teachers who were exemplary or needs improvement) - For teachers identified as needing improvement, IMs will build in specific supports (i.e., Plan of Improvement, coaching, observing exemplary teachers) • Increased responsibility <ul style="list-style-type: none"> - Exemplary teachers are eligible to be a lead teacher, mentor, or coach - Exemplary teachers are also invited to model/share best practices and lead Professional Learning Communities (PLC) (for which they receive a stipend) • Tenure <ul style="list-style-type: none"> - Non-tenured teachers are offered a contract renewal for the next year based on performance. In order to be eligible for tenure, non-tenured teachers must be offered a contract renewal for the next year - For tenured teachers at this time, performance cannot affect tenure status, as it is regulated by Connecticut state law <ul style="list-style-type: none"> » However, tenured teachers who receive a rating of "developing" for more than two consecutive years can move down to "needs improvement" (at the IM's discretion) • Dismissal <ul style="list-style-type: none"> - Close to the end of the year, district officials meet with principals and IMs to identify teachers who are rated "needs improvement" and do not show improvement. The teachers' union also has conversations with these teachers. The assistant superintendent makes final recommendations on dismissal. - Thus far, TEVAL has not resulted in a dismissal, although the district is committed to terminating teachers if necessary - In 2010–11, <ul style="list-style-type: none"> » 75 teachers had scores that put their jobs at risk » 34 of those teachers voluntarily resigned, including 16 tenured teachers » 15 teachers jobs were preserved due to technical issues » The rest improved enough to stay. There were no terminations

<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<ul style="list-style-type: none"> • If a teacher disagrees with the IM's rating, he/she may attach a statement to the evaluation. These issues are often resolved by a meeting with the teacher and the IM, facilitated by the TEVAL coordinator • If a teacher evaluated by a TPV, meaning that the teacher is exemplary or needs improvement, disagrees with the IM or the TPV's ratings, the reports of the IM and TPV are compared. If the reports agree, the rating stands. If the reports disagree, the case is examined by the assistant superintendent in collaboration with the president of the New Haven Federation of Teachers (NHFT)
<p><i>If performance results are available, what is the score distribution?</i></p>	<p>Distribution of ratings from 2010-2011 school year:</p> <ul style="list-style-type: none"> • Exemplary: 8% • Strong: 38% • Effective: 28% • Developing: 9% • Needs Improvement: 3% • No ratings: 11% • Other ratings (for retired teachers, teachers on leave, etc): 3%
<p>TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT</p>	
<p><i>What role did teachers play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • Six teachers and six administrators sat on the TEVAL committee to design the new teacher evaluation system and a much larger teacher working group was consulted during the design process • A TEVAL working group, open to any interested teacher, had input into the evaluation process and rubric design
<p><i>What role did the union play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • As part of the district's <u>School Change Initiative</u>, NHPS partnered with the New Haven Federation of Teachers (NHFT) in October 2009 to make recommendations to the Board of Education to reform the evaluation system • The 2009 teachers' contract agreed to a new teacher-evaluation system • The union president also sat on the committee to reform the principal evaluation system. TEVAL was part of broader conversations around school reforms in the district • Both the union and the district are contacted if there is a discrepancy between an evaluator and a teacher
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<ul style="list-style-type: none"> • The TEVAL working group of teachers continues to meet to discuss revisions • <u>School Climate Surveys</u>, completed annually by teachers at each school, also include questions about TEVAL

SITE PROFILES

GLOSSARY	
<i>Growth goal</i>	<ul style="list-style-type: none"> • Set by teachers and instructional managers • Minimum of two and maximum of four required for each teacher • Based on CMT data or other approved assessments
<i>Instructional Manager (IM)</i>	Principal, assistant principal or other administrative leader accountable for teacher's evaluation and development
<i>Third-Party Validator (TPV)</i>	<ul style="list-style-type: none"> • Former teachers who are not currently affiliated with district but have demonstrated effectiveness in the classroom in the past • Hired with a joint contract, selected by schools in consultation with district and union

TIMELINE



SOURCES

Connecticut State Department of Education. *Bureau of Student Assessment*. Retrieved from <http://www.csde.state.ct.us/public/cedar/assessment/index.htm>

New Haven Public Schools. *Demographics*. Retrieved from <http://www.nhps.net/nhpsdemographics>

New Haven Public Schools. (2010). *NHPS teacher evaluation and development process: Teachers and administrators guide*. Retrieved from [http://www.nhps.net/sites/default/files/2_TEVAL - Teachers and Administrators Guide 100825.pdf](http://www.nhps.net/sites/default/files/2_TEVAL_-_Teachers_and_Administrators_Guide_100825.pdf)

New Haven Public Schools. (2010). *Teacher evaluation and development: Introduction to the process*. Retrieved from [http://nhps.net/sites/default/files/1_NHPS TEVALDEV Introduction - Aug 20103.ppt](http://nhps.net/sites/default/files/1_NHPS_TEVALDEV_Introduction_-_Aug_20103.ppt)

New Haven Public Schools. (2011). *Innovative contracts: breaking the mold*. American Federation of Teachers. Retrieved from http://www.aft.org/pdfs/teachers/teach11materials/t145_innovativepres.pdf

Sawchuk, Stephen. (2011, November 15). Partnership yields revised evaluation system in New Haven. *Education Week*. Retrieved from <http://www.edweek.org/ew/articles/2011/11/16/12collab-newhaven.h31.html>

PITTSBURGH, PA

PITTSBURGH PUBLIC SCHOOLS' PROFESSIONAL GROWTH SYSTEM

OVERVIEW

Pittsburgh Public Schools' (PPS) Professional Growth System is part of a larger reform effort in Pittsburgh known as *Empowering Effective Teachers*, aimed at improving the recruitment, placement, evaluation, retention, and support of teachers. In PPS, teacher evaluation is a tool for improving teacher quality through development and growth while increasing accountability for student outcomes. According to the district, the impetus for the change came from teachers, who recognized the need for a better system for evaluation and feedback, leading to the implementation of a new observation system district-wide in 2010–11. PPS has received nearly \$90 million in private and public funds to use toward its teacher effectiveness reforms over six and half years, including \$40 million from the Bill & Melinda Gates Foundation, with whom it works as an “intensive partnerships for effective teaching” site.

PPS collects three measures of teacher effectiveness: teacher practice, student learning and growth, and other student outcomes. Teacher practice considers a teacher’s planning and preparation, classroom environment, professional responsibilities, and teaching and learning, based largely on observations conducted by administrators. Most teachers receive at least four observations a year (two formal and two informal), while non-tenured teachers receive at least eight (four formal and four informal). PPS is also measuring teachers’ contributions to student learning and growth using value-added data and administering the *Tripod student survey* as a measure of other student outcomes. Currently, ratings in PPS’s teacher evaluation system only consider teacher practice based on observation. PPS plans to include student-level learning and growth in teachers’ evaluation ratings beginning in 2013–14, and is considering how to include Tripod survey results and school-level growth scores in the future.

A standout feature of Pittsburgh’s teacher evaluation system and its development has been the implementation of the system in phases. PPS has piloted each portion for its evaluation system and made revisions as needed before including the results in teachers’ final ratings. District leaders chose this approach to ensure that as many stakeholders as possible understand and value each piece, and that each measure is methodologically sound and accurately measures teacher effectiveness. Pittsburgh Public Schools’ Professional Growth System is also unique in that it is separate from Employee Improvement Plans, ensuring that teacher development remains the primary purpose as the system is developed.

Looking forward, PPS faces several challenges related to teacher evaluation, including achieving consistent engagement and communication at all levels, building confidence in the rigor and fairness of its evaluation system, and maximizing the information produced through evaluation to facilitate professional growth and dramatically improve student outcomes.

ABOUT THE SCHOOL SYSTEM

<i>Number of students</i>	25,000
<i>Number of schools</i>	60
<i>Number of teachers</i>	2,000

BACKGROUND	
<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	Yes, 24 schools in 2009–10
<i>In effect since</i>	2010–11 (including only the teacher practice portion of the evaluation system)
<i>Are there plans for additional phases/components</i>	<p>Yes, the district will include student outcomes in its evaluation system:</p> <ul style="list-style-type: none"> • Teachers currently receive value-added scores and student survey results, but these data are used primarily for formative, not evaluative, purposes • Student learning and growth results will be included in evaluation scores beginning in 2013–14 • Use of student survey results and school-level growth in evaluation score to be determined
<i>Who gets evaluated?</i>	<ul style="list-style-type: none"> • All teachers, except those on an Employee Improvement Plan, which is a separate process • In 2014–15, PPS plans to update the principal evaluation system to align with the teacher evaluation and professional growth system
<i>How often do teachers receive a rating?</i>	<ul style="list-style-type: none"> • All teachers evaluated through the district’s <u>Research-based, Inclusive System of Evaluation (RISE) rubric</u> receive a rating every year based on 12 of the rubric’s 24 components • Currently, non-tenured teachers receive ratings twice a year • Teachers on an Employee Improvement Plan are evaluated through a different system
<i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i>	Not applicable
STUDENT ACHIEVEMENT MEASURES	
<i>Assessments to measure student achievement</i>	<ul style="list-style-type: none"> • Pennsylvania System of Standardized Assessment (PSSA) <ul style="list-style-type: none"> - Reading: grades 3–8 and 11 - Math: grades 3–8 and 11 - Writing: grades 5, 8, and 11 - Science: grades 4 and 8 • Nationally normed assessments (e.g., PSAT Reading, PSAT Writing, PSAT Math) • Locally developed curriculum-based assessments (CBAs) <ul style="list-style-type: none"> - Math (including Algebra and Geometry): grades 6–11 - English (including African-American literature): grades 6–12 - Science (including earth science, life science, biology, chemistry, and physics): grades 6–11 - History (including civics, world history, and U.S. history): grades 8–11
<i>Are any new assessments in development to use with teacher evaluations?</i>	Over time, PPS plans to expand and improve its assessment portfolio to include more grades and subjects.

SITE PROFILES

<p><i>What academic/achievement criteria are included?</i></p>	<p>In the 2011–12 school year, academic/achievement criteria are not a factor in teachers’ evaluation rating, but PPS plans to include student learning and growth measures based on value-added data beginning in 2013-14.</p> <ul style="list-style-type: none"> • Value-added data are currently available for approximately 40 percent of PPS teachers, and shared with teachers for formative purposes • Initially, student learning and growth will be used only for teachers with value-added data • PPS is also considering including a school-level growth measure as a portion of evaluation starting in 2013–14 as PPS continues to work to align with state policy
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<ul style="list-style-type: none"> • Exact weighting has not yet been determined, but PPS’s goal is to align the weights with a state evaluation system that is in development, and has proposed 50% teacher practice and 50% student outcomes • Over time, PPS plans to expand and improve its assessment portfolio to include more grades and subjects
<h3>NONACADEMIC MEASURES</h3>	
<p><i>What nonacademic criteria are included?</i></p>	<ul style="list-style-type: none"> • “Teacher practice” includes 24 components in four areas based on Charlotte Danielson’s Framework for Teaching: <ul style="list-style-type: none"> - Planning and preparation - Classroom environment - Teaching and learning - Professional responsibilities • In the future, evaluation may include a measure of a teacher’s contribution to other student outcomes, including students’ input on their experience in the classroom <ul style="list-style-type: none"> - PPS is administering the Tripod student survey in the vast majority of classrooms this year - The Tripod survey measures students’ classroom experience, and the Measures of Effective Teaching (MET) Project’s research has shown that Tripod survey results have a positive correlation to student value-added results - Currently, this tool is only used for growth and formative feedback
<p><i>How much does the achievement component count in a teacher’s final rating?</i></p>	<ul style="list-style-type: none"> • 2011–12: Teacher’s final rating based entirely on teacher practice • 2013–14: Exact weighting not yet determined, but will include teacher observation and student learning and growth
<h3>CLASSROOM OBSERVATIONS</h3>	
<p><i>Who are the observers who rate teachers’ effectiveness?</i></p>	<ul style="list-style-type: none"> • Principals and assistant principals are the primary observers, but they can use the assistance of subject-area experts in the observation and conference processes • At the high school level, at least some observations are conducted by trained experts in that subject area • Starting in 2012–13, PPS will have a new leadership role for effective teachers called Instructional Teacher Leader 2 (ITL2) <ul style="list-style-type: none"> - ITL2 is a three-year role that includes conducting teacher observations - In the first year, ITL2s will contribute to formative feedback only - In years 2 and 3, ITL2s will contribute to teachers’ summative evaluations

<p><i>What training do observers receive?</i></p>	<ul style="list-style-type: none"> • There were several days of initial evaluation training for leadership teams the summer before the district’s teacher observation system (RISE) was implemented • All district evaluators participate in the Instructional Quality Assurance and Certification (IQA-C) Process, which builds ongoing training provided to observers. There are two levels to the IQA-C Process, focusing on accuracy and consistency in scoring, and instructional feedback and support: <ul style="list-style-type: none"> - Level 1: ensure ability to accurately identify evidence of effective teaching and accurately score evidence against the RISE rubric <ul style="list-style-type: none"> » Currently, more than 90 percent of principals and Career Ladder teachers are certified Level 1 observers - Level 2: ensure quality of instructional feedback and support • District leadership from the Pittsburgh Federation of Teachers, assistant superintendents, and consultants provide additional training days during the year. RISE District/PFT Leadership Team has completed more than 270 on-site visitations.
<p><i>Are observers compensated for this work?</i></p>	<ul style="list-style-type: none"> • Instructional Teacher Leaders 2 will receive an additional \$9,300 on top of their regular teaching salary when they begin this work next year
<p><i>How often are teachers observed?</i></p>	<ul style="list-style-type: none"> • Non-tenured teacher: 8+ observations/year (4 formal and 4 informal) • Teachers in career ladder roles (promotional roles based on evidence of high effectiveness): 4+ observations/year (2 formal and 2 informal) • Experienced (tenured) teachers: <ul style="list-style-type: none"> - Two-thirds of experienced teachers: 4+ observations/year (2 formal and 2 informal) - One-third of experienced teachers: participate in a year of “Supported Growth” <ul style="list-style-type: none"> » Intended to improve teacher practice » These teachers focus on one or more of the RISE components » They work in a peer cohort to monitor and present progress » Administrators provide support and feedback only » There are no formal observations » Final rating based on averaged rating by peers and principal » For 2011–12, the teachers participating in Supported Growth were selected from RISE school-based leadership teams, teacher volunteers, and lottery when necessary. PPS is still developing the selection process for Supported Growth moving forward • Formal observations must last longer than 30 minutes and include evidence-sharing by the principal, teacher self-assessment, and a post-observation conference • Informal observations generally last 10 to 15 minutes
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<ul style="list-style-type: none"> • New teachers: At least two formal observations in the first semester • Experienced teachers: By end of year
<p><i>Are observations announced or unannounced?</i></p>	<p>One observation must be announced.</p>
<p><i>What is the basis for the observation rubric?</i></p>	<p>PPS’s observation rubric is based on Charlotte Danielson’s Framework for Teaching.</p>

SITE PROFILES

<p><i>What is included in the observation scoring rubric?</i></p>	<ul style="list-style-type: none"> • Charlotte Danielson’s Framework for Teaching has 24 components across four domains: <ul style="list-style-type: none"> - Planning and preparation - Classroom environment - Professional responsibilities - Teaching and learning • 12 “Power Components” are used for the rating <p>See a copy of the rubric here.</p>
<p><i>How are observers held accountable for their scoring?</i></p>	<p>Through the IQA-C Process (mentioned above), PPS is continuously strengthening the accuracy, validity, and reliability of evaluators to collect evidence during observations and score this evidence against the rubric.</p> <ul style="list-style-type: none"> • Currently, more than 90 percent of principals and Career Ladder teachers are certified Level 1 observers through IQA-C • An online software tool now supports the observation process • On-site training and support offered during monthly visits from district staff, supported by a consultant from the Danielson Group
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • Yes, after formal observations • Teachers also receive electronic feedback from informal observations and can request a debrief directly
<h3>ACCURACY, VALIDITY, AND RELIABILITY</h3>	
<p><i>Is there a process to validate non-standardized assessments?</i></p>	<p>Not applicable</p>
<p><i>Are there procedures to ensure that observation scores are valid?</i></p>	<p>In addition to the IQA-C Process and other training described above, PPS helped pilot the Validation Engine Project through the Measures of Effective Teaching Project (MET)</p>
<p><i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i></p>	<p>This summer, PPS will work with experts to begin the process of looking at the relationships among the multiple measures now in place</p>
<p><i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i></p>	<p>Pittsburgh is closely following the results of national research efforts underway to develop the best approaches to measuring and improving teacher effectiveness, including the Measures of Effective Teaching (MET), which is exploring each measure now in place in PPS</p>
<h3>USING AND REPORTING EVALUATION RESULTS</h3>	
<p><i>What are the rating categories?</i></p>	<ul style="list-style-type: none"> • There are four rating categories: <ul style="list-style-type: none"> - Unsatisfactory - Basic - Proficient - Distinguished • Currently, teachers’ ratings come from their score on 12 of the 24 “Components of Practice” across four “domains” of effective teaching consistent with the Charlotte Danielson Framework for Teaching

<p><i>How are results reported?</i></p>	<p>Teachers receive their ratings, but results are not reported to the public.</p>
<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<p>Although multiple measures will not be used for teachers' end-of-year ratings until 2013–14, they already affect professional development, pay, opportunities for increased responsibility, and tenure</p> <ul style="list-style-type: none"> • Professional development/improvement plan <ul style="list-style-type: none"> - In a given year, a third of experienced teachers will participate in a year of supported growth, where they focus on one or more of the observation rubric components, but are not formally observed • Promotion/pay <ul style="list-style-type: none"> - Significant awards in place at the district, school, and team levels based on contribution to student learning and growth based on value-added scores - A new salary schedule links career earnings directly to evidence of effectiveness based on multiple measures for teachers hired after June 2010 • Increased responsibility (e.g., mentor teacher, etc.) <ul style="list-style-type: none"> - PPS has developed six career ladder roles that seek to place effective teachers with the highest need students <ul style="list-style-type: none"> » Learning Environment Specialist (LES) » Clinical Resident Instructor (CRI) » Promise-Readiness Corps (PRC) » Promise-Readiness Corps/Clinical Residency Instructor » K–8 Instructional Teacher Leaders 2 (ITL2) (to be launched in 2012–13) » Secondary Instructional Teacher Leaders 2 (ITL2) (to be launched in 2012–13) • Tenure <ul style="list-style-type: none"> - Pre-tenured teachers must receive six satisfactory ratings before they receive tenure
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<p>The district and union have appeal processes</p>
<p><i>If performance results are available, what is the score distribution?</i></p>	<p>Results are not yet available</p>
<p>TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT</p>	
<p><i>What role did teachers play in developing the evaluation system?</i></p>	<p>Teachers were directly involved in the development of RISE observation system in three primary ways:</p> <ol style="list-style-type: none"> 1. Fleshing out the details of the observation system. Representatives from each pilot school (the principal and one lead teacher), along with the Pittsburgh Federation of Teachers (PFT) and district leaders, formed a design team to develop the details of the evaluation program. At this point, more than 400 teachers have been involved in the design of the measures 2. Continuous refinement. In 2009, PPS formed leadership teams consisting of four teachers and a principal from each school. These teams continue to be responsible for leading training and disseminating information on multiple measures at the school level, monitoring their implementation, and offering input on the refinement of the measures, but not for conducting observations 3. Union involvement. In addition, the PFT played a large and decisive role developing RISE (see below)

SITE PROFILES

<p><i>What role did the union play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • The union served as a collaborative partner with the district to create RISE. Two high-level staff members from both the Pittsburgh Federation of Teachers and Pittsburgh Public Schools formed the PFT/PPS RISE Leadership Team responsible for creating RISE • Representatives from the PFT have also participated in the development and training process for value-added measures • The partnership between PPS and the PFT is documented in more detail in <i>Forging a New Partnership: The Story of Teacher Union and School District Collaboration in Pittsburgh</i>
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<p>Yes, the system is continually refined, and the RISE design team contributed to the decision-making around this process.</p>
<h3>GLOSSARY</h3>	
<p><i>Empowering Effective Teachers</i></p>	<ul style="list-style-type: none"> • Pittsburgh’s comprehensive plan to increase teacher effectiveness • Funded by the Bill & Melinda Gates Foundation, the Teacher Incentive Fund (TIF), and other private and public donors • Goals include: <ul style="list-style-type: none"> - Increase the number of highly effective teachers - Increase the exposure of high-need students to highly effective teachers - Ensure that all teachers work in learning environments that support their ability to be highly effective
<p><i>Tripod student survey</i></p>	<ul style="list-style-type: none"> • Student survey designed to measure school quality, teaching effectiveness, and student engagement
<p><i>Instructional Quality Assurance and Certification (IQA-C) Process</i></p>	<ul style="list-style-type: none"> • Ongoing training for all district observers • There are two levels to the IQA-C Process, focusing on accuracy and consistency in scoring, and instructional feedback and support: <ul style="list-style-type: none"> - Level 1: ensure ability to accurately identify evidence of effective teaching and accurately score evidence against the <u>RISE rubric</u>. - Level 2: ensure quality of instructional feedback and support
<p><i>Instructional Teacher Leader 2 (ITL2)</i></p>	<ul style="list-style-type: none"> • New leadership role to go into effect in 2012–13 • ITL2 is a three-year role that includes conducting teacher observations • In the first year, ITL2s will contribute to formative feedback only • In years 2 and 3, ITL2s will contribute to teachers’ summative evaluations
<p><i>Research-based Inclusive System of Education (RISE)</i></p>	<ul style="list-style-type: none"> • The PPS teacher observation system
<p><i>Supported growth</i></p>	<ul style="list-style-type: none"> • Year in which experienced teachers work in peer cohorts to focus on one or more of the RISE components • Teachers do not have any formal observations in this year

TIMELINE

Not available.

SOURCES

Hamill, S. (2011). *Forging a new partnership: The story of teacher union and school district collaboration in Pittsburgh*. New York: The Aspen Institute Education & Society Program. Retrieved from <http://www.aspeninstitute.org/sites/default/files/content/docs/education/ed-PittRprt.pdf>

Pittsburgh Public Schools. *Empowering Pittsburgh Teachers in Pittsburgh Public Schools*. Education Committee Update January 10, 2012.

Pittsburgh Public Schools. *The Empowering Effective Teachers Plan*. Retrieved from <http://www.pps.k12.pa.us/170220210151928197/lib/170220210151928197/RISE-Implementation-Overview.pdf>

WASHINGTON, DC

IMPACT

OVERVIEW

When Michelle Rhee became chancellor of the District of Columbia Public Schools (DCPS) in 2007, she made it one of her top priorities to revamp the district’s teacher evaluation system as a central strategy to accelerate student achievement. Within two years, DC IMPACT went into effect with the goal of drastically improving the quality of DCPS teachers.

IMPACT includes five components: 1) teacher-level student achievement; 2) school-level student achievement; 3) teacher performance in the classroom using the Teaching and Learning Framework; 4) teachers’ commitment to the school community; and 5) professionalism. The weight of each component and the measure of student growth varies based on the availability of data. For example, teachers of tested grades and subjects receive an individual value-added score that counts for half of their final rating. Meanwhile, for teachers of untested grades and subjects for whom value-added data are unavailable, student achievement is measured using teacher-assessed student achievement data. For these teachers, the student achievement component is weighted much less in the final rating—just 10 percent—and other components, including teacher performance in the classroom, count more heavily.

One of IMPACT’s most distinctive features is its use of master educators whose full-time job is to observe teachers as impartial evaluators. Master educators are content specialists who conduct two of teachers’ five required observations each year.

Several rewards and consequences for teachers are also tied to IMPACT. Teachers who are rated “highly effective” have the potential to earn more than \$130,000 per year (compared with \$87,500 previously), and to reach the maximum salary in less than 10 years (compared with 21 years previously). In return, however, teachers must also opt to forgo a number of job protections, including some options if a teacher is excessed because their position is eliminated and they cannot find a position in another school. Approximately 70 percent of teachers eligible for a bonus and 80 percent of teachers eligible for a salary increase have accepted the offer. These reforms are supported partly through private funding.

DCPS has faced several challenges related to IMPACT. DCPS believes that principals, as instructional leaders, ought to spend a large portion of their time observing and providing feedback to teachers to ensure that students are receiving high-quality instruction every day. At the same time, the district recognizes that principals have many other responsibilities and that their role in IMPACT is very time-consuming. The district is therefore considering ways to streamline the system to reduce the workload for principals without compromising its integrity.

The student achievement measures represent another challenge. While DCPS believes that its value-added measure is the fairest and most accurate way to measure teachers’ impact on student performance, it also recognizes that the measure presents certain challenges—including that it does not reflect everything a teacher teaches, many teachers find it difficult to understand, and data aren’t available until after the school year ends. For teachers for whom value-added data are not available, the district has found it challenging to ensure the rigor and consistency of alternative measures.

ABOUT THE SCHOOL SYSTEM

<i>Number of students</i>	45,000
<i>Number of schools</i>	125
<i>Number of teachers</i>	4,000

BACKGROUND	
<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	No
<i>In effect since</i>	2009–10
<i>Are there plans for additional phases/components?</i>	No
<i>Who gets evaluated?</i>	<ul style="list-style-type: none"> All school-based personnel, including teachers, student support professionals, library media specialists, counselors, school-based social workers, school-based psychologists, related service providers, special education coordinators, program coordinators and deans, instructional coaches, educational aids, office staff, and custodial staff <u>Principals are also assessed using an aligned system</u>
<i>How often do teachers receive a rating?</i>	Once a year
<i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i>	Not applicable
STUDENT ACHIEVEMENT MEASURES	
<i>Assessments to measure student achievement</i>	<ul style="list-style-type: none"> District of Columbia Comprehensive Assessment System (DC-CAS) <ul style="list-style-type: none"> - Reading: grades 3–8 and 10 - Math: grades 3–8 and 10 - Composition: grades 4, 7, and 10 - Science: grades 5 and 8 - Biology: students taking biology in grades 9–12 Administrator-approved assessments of student growth over the year (for teachers of grades and subjects for which DC-CAS data are not available)
<i>Are any new assessments in development to use with teacher evaluations?</i>	DCPS plans to implement new assessments to provide value-added data for more grades and courses, including K–2, and secondary English, math, science, and social studies.

SITE PROFILES

<p><i>What academic/achievement criteria are included?</i></p>	<p>DC IMPACT includes three possible student achievement measures: a school value-added score whenever available, an individual value-added score (for teachers of tested grades and subjects), or a growth or mastery score based on teacher-assessed student achievement (for teachers of untested grades and subjects):</p> <ul style="list-style-type: none"> • 1) School value-added (SVA) score (see p. 46 of guidebook) <ul style="list-style-type: none"> - Difference between a school’s “likely score” on DC-CAS predicted by value-added model and a school’s actual student performance • 2) Individual value-added (IVA) score <ul style="list-style-type: none"> - Difference between a class’s “likely score” on DC-CAS predicted by value-added model and a class’s actual score - Used for teachers of reading or math in grades 4–8 only (grades for which DC-CAS data are available from the previous year) • 3) <u>Teacher-assessed student achievement data (TAS)</u> <ul style="list-style-type: none"> - TAS is a measure of student learning over the course of the year as measured by assessments other than the DC-CAS used for teachers of untested grades and subjects only to measure student learning during the year - Teachers of untested grades and subjects choose the assessment(s) they will use for this component, as well as weights assigned to each assessment and an achievement target. School administrators must approve all assessments and weights, and the district provides recommendations for appropriate assessments and weights <ul style="list-style-type: none"> » For example, first-grade reading: 70%+ students increase at least one proficiency level or are at least proficient on the Text Reading and Comprehension Assessment » For example, secondary math: Average class score of 70%+ on teacher-created or off-the-shelf assessment that measures student mastery of DCPS standards - Teachers meet with their school administrators in the fall to get approval of assessments, weights, and targets. An operations team in the district office then audits all goals - In the spring, teachers present their student achievement data to their school administrator, who verifies them and assigns each teacher a score based on a rubric that provides descriptions of four performance levels (see scoring rubric on p. 38 of guidebook)
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<p>The weight of the student achievement components varies based on data available for a teacher:</p> <ul style="list-style-type: none"> • Teachers with individual value-added data: <ul style="list-style-type: none"> - School value-added: 5% - Individual value-added: 50% • Teachers without individual value-added data: <ul style="list-style-type: none"> - School value-added: 5% (Note: if SVA not available, “commitment to the school community,” a nonacademic component, will count for an additional 5% of final score) - Teacher-assessed student achievement data: 10%

NONACADEMIC MEASURES

<p><i>What nonacademic criteria are included?</i></p>	<p>There are three nonacademic measures included in IMPACT:</p> <ol style="list-style-type: none"> 1. Teaching and Learning Framework that <ul style="list-style-type: none"> - assesses teacher practice in the classroom (rubric on pp. 14–35 of guidebook) - DCPS consulted 20 sources to create the Teaching and Learning Framework, but ultimately created its own unique framework 2. Commitment to the School Community, measures a teacher’s work as a member of a school community (rubric on pp. 40–45 of guidebook) 3. Core Professionalism (rubric on pp. 53–55 of guidebook) <ul style="list-style-type: none"> - Includes following school policies and procedures, interacting with school and community members in a respectful manner - Assessed twice per year
<p><i>How much do the nonacademic measures count in a teacher’s final rating?</i></p>	<p>Teachers receive a final evaluation score between 100 and 400. To calculate the score, IMPACT multiplies the component score (1–4) by the weight of each component. The weight of the nonacademic criteria varies based on data availability:</p> <ul style="list-style-type: none"> • Teachers with individual value-added data: <ul style="list-style-type: none"> - Teaching and Learning Framework: 35% - Commitment to the School Community: 10% - Core Professionalism <ul style="list-style-type: none"> » If teacher “meets standard” for all parts, no impact on final score » If teacher is “slightly below standard” on any component in either assessment cycle, 10 points will be deducted from final IMPACT score (with an additional 10 points subtracted if the teacher receives a “slightly below standard” in both assessment cycles) » if teacher is “significantly below standard” on any part, 20 points will be deducted from final IMPACT score (with an additional 20 points subtracted if the teacher receives a “significantly below standard” in both assessment cycles) • Teachers without individual value-added data: <ul style="list-style-type: none"> - Teaching and Learning Framework: 75% - Commitment to the School Community: 10% (Note: if school value-added score is not available, commitment to school community counts for 15% of final score) - Core Professionalism (see above for explanation of scoring)

SITE PROFILES

CLASSROOM OBSERVATIONS	
<p><i>Who are the observers who rate teachers' effectiveness?</i></p>	<p>There are two kinds of observers:</p> <ul style="list-style-type: none"> • Teachers' school administrators • Master educators (MEs) <ul style="list-style-type: none"> - MEs are expert teachers whose full-time job is to travel from school to school as impartial evaluators - MEs are content specialists in the following areas: <ul style="list-style-type: none"> » Bilingual education » Early childhood education » Electives (including music, physical education, art, and foreign language) » Elementary (upper and lower elementary) » Secondary (e.g., math, science, English, social studies) » Special education • Approximately 40 master educators work in IMPACT Approximately 2,400 people have applied for the position in the first three years (2009–10 to 2011–12) • For more on MEs, including selection process and details on training and support, see IMPACT's website and slides from an April 2011 conference
<p><i>What training do observers receive?</i></p>	<ul style="list-style-type: none"> • Master educators (MEs): <ul style="list-style-type: none"> - Six-week summer institute, including: <ul style="list-style-type: none"> » Three weeks of rubric analysis, practicing consistency in scoring by observing teaching videos (video norming), report writing, and post-observation conference practice » Two weeks of practice observations in summer-school classrooms » One week of logistics training and teacher resource development - MEs also meet biweekly throughout the year, as a whole group and in content-specific cohorts, to engage in additional professional development • Administrators also receive training, which consists of sessions during the Summer Leadership Academy and additional follow-up during monthly Principals' Academies. Training includes: <ul style="list-style-type: none"> - Practicing consistency in scoring by observing teaching videos (video norming) - Informal observations in classrooms with master educators and instructional superintendents - Specific support on written feedback and goal-setting/tracking data for the Teacher Assessed Student achievement component • As more administrators become familiar with IMPACT, there has been less training • DCPS is in the process of developing online trainings and calibration tools
<p><i>Are observers compensated for this work?</i></p>	<ul style="list-style-type: none"> • School administrators do not receive additional compensation for this work • A master educators' entire job is to conduct this work, and they are compensated through their salary

<p><i>How often are teachers observed?</i></p>	<ul style="list-style-type: none"> • Most teachers: Five formal observations/year (30 minutes or longer) <ul style="list-style-type: none"> - Administrators conduct three observations - Master educators conduct two observations • Highly effective teachers: <ul style="list-style-type: none"> - Teachers earning a “highly effective” rating for the previous two years will receive two observations (one from administrator and one from master educator) by December 1 - If they receive an average score of 3.5 or higher out of 4 on the observation rubric, they can waive the last three observations for the year
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<ul style="list-style-type: none"> • Administrator observations: <ul style="list-style-type: none"> - Observation 1: September 12–December 1 - Observation 2: December 1–March 1 - Observation 3: March 1–June 1 • Master educator observations: <ul style="list-style-type: none"> - Observation 1: September 12–February 1 - Observation 2: February 1–June 1
<p><i>Are observations announced or unannounced?</i></p>	<ul style="list-style-type: none"> • The first administrator observation is announced • All other observations are unannounced
<p><i>What is the basis for the observation rubric?</i></p>	<p>DCPS created its own performance rubric to use with DC IMPACT, but it is influenced by more than a dozen frameworks, including Charlotte Danielson’s <i>Framework for Teaching</i>, Martin Haberman’s <i>Star Teacher, Teach for America’s Teaching as Leadership</i>, and Connecticut’s <i>Common Core of Teaching</i>, among others</p>
<p><i>What is included in the observation scoring rubric?</i></p>	<ul style="list-style-type: none"> • DC IMPACT uses the <u>Teaching and Learning Framework</u>, which has three domains: <ul style="list-style-type: none"> - Plan - Teach - Increase effectiveness • In the 2011–12 school year, however, teachers will be assessed only on the “Teach” domain
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • Master educators are evaluated regularly based on: <ul style="list-style-type: none"> - Rating accuracy - Written reports - Post-observation conferences - Facilitation of professional development - Communication and collaboration - Dependability and professionalism • Evaluations conducted through: <ul style="list-style-type: none"> - “Ride-alongs,” in which senior MEds attend observations and conferences with the MEds they manage to provide feedback and assess ME performance - Analysis of written work - Observations of professional development - Data review • Using multiple evaluators ensures more accurate scores and additional feedback for teachers • DCPS checks that ME and administrator ratings are correlated

SITE PROFILES

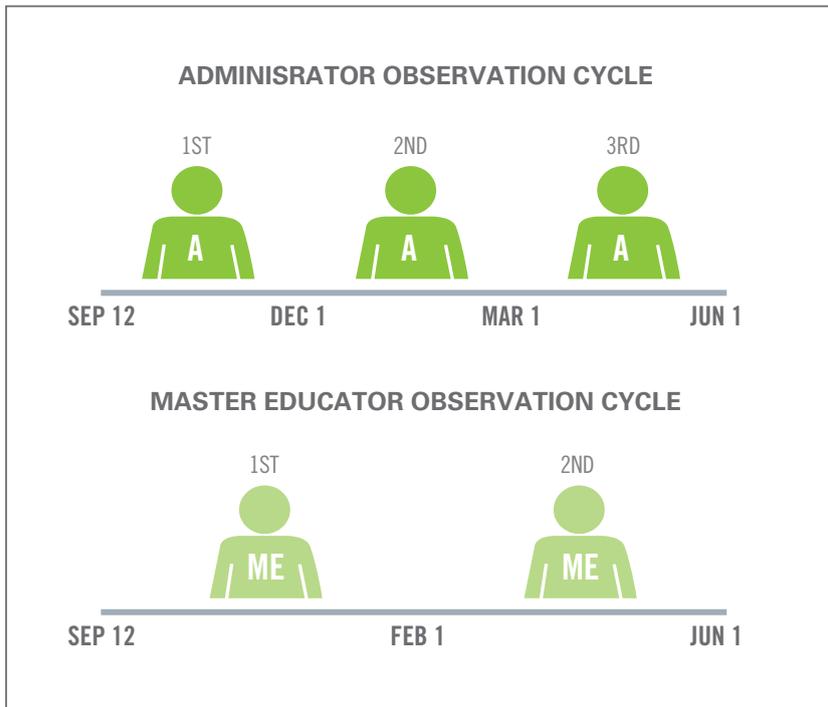
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • Yes, observers must meet with teachers within 15 calendar days of the observation • Teachers also receive written feedback from all formal observations through a web-based portal
<p>ACCURACY, VALIDITY, AND RELIABILITY</p>	
<p><i>Is there a process to validate non-standardized assessments?</i></p>	<ul style="list-style-type: none"> • A DCPS operation team audits goals for the teacher-assessed student achievement data component • Principals also approve teacher goals and assessments, and validate achievement data at the end of the year
<p><i>Are there procedures to ensure that observation scores are valid?</i></p>	<ul style="list-style-type: none"> • Using multiple evaluators ensures more accurate scores and additional feedback for teachers • Master educators participate in ongoing training to ensure that they are assigning consistent scores during observations • Master educators and principals also conduct informal observations together to ensure rating consistency
<p><i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i></p>	<p>DCPS has found that the correlation between student value-added data and observations is about 0.4, signifying a fairly strong relationship between the measures.</p>
<p><i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i></p>	<p>DCPS is looking at the correlation between academic and nonacademic measures, but has not yet released those data publicly.</p>
<p>USING AND REPORTING EVALUATION RESULTS</p>	
<p><i>What are the rating categories?</i></p>	<p>There are four rating categories:</p> <ul style="list-style-type: none"> • Highly effective • Effective • Minimally effective • Ineffective
<p><i>How are results reported?</i></p>	<ul style="list-style-type: none"> • Results are reported to teachers and administrators • Distribution of scores across DCPS are reported publicly

<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<p>Yes, evaluation results are tied to decisions regarding professional development, promotion, pay, and dismissal.</p> <ul style="list-style-type: none"> • Professional development <ul style="list-style-type: none"> - Results are used to inform both structured learning cycles that instructional coaches facilitate, and district-wide professional development • Promotion/Pay <ul style="list-style-type: none"> - Teachers have the opportunity to opt into salary system where they can move up the pay scale more quickly, reaching the maximum salary in nine years, compared with 21 years under the traditional system. In exchange they must forgo some work security, including some options if a teacher is excessed and cannot find a position in another school. <ul style="list-style-type: none"> » Teachers earning a rating of “highly effective” are eligible for bonuses of up to \$25,000 » Teachers earning a rating of “highly effective” for at least two consecutive years are eligible for a base salary increase of up to \$27,000 • Dismissal <ul style="list-style-type: none"> - Teachers earning an “ineffective” rating are subject to dismissal - Teachers earning a “minimally effective” rating for two consecutive years are also subject to dismissal • Other <ul style="list-style-type: none"> - IMPACT ratings are considered for other awards/recognitions, including Rubenstein Awards for Highly Effective Teaching and A Standing Ovation for DC Teachers gala
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<ul style="list-style-type: none"> • Through the chancellor’s appeals process, DCPS provides teachers who earn ineffective or minimally effective ratings an opportunity to submit an appeal if they do not believe the evaluation procedures were followed appropriately (e.g., a post-observation conference was not provided within 15 days). Teachers may appeal only the evaluation process, not the observation scores themselves • The Washington Teachers’ Union also provides teachers with an opportunity to submit appeals
<p><i>If performance results are available, what is the score distribution?</i></p>	<p>In 2010–11:</p> <ul style="list-style-type: none"> • Highly effective: 16% • Effective: 67% • Minimally effective: 13% • Ineffective: 2% (these teachers were dismissed) • Ineligible for score: 2%
<p>TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT</p>	
<p><i>What role did teachers play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • In IMPACT’s first year, DCPS hosted 50 feedback sessions, allowing approximately 500 teachers to provide feedback on the design • In IMPACT’s second year, DCPS held more than 100 feedback sessions and focus groups with more than 1,000 educators to discuss revisions to IMPACT and consider various policy approaches • DCPS continues to engage teachers through in-person focus groups, online surveys, and other forums

SITE PROFILES

<p><i>What role did the teachers union play in developing the evaluation system?</i></p>	<p>According to DCPS, engaging the Washington Teachers' Union (WTU) was a priority for Chancellor Rhee. DCPS was in constant conversation with the union leadership during the IMPACT design process, and asked the WTU and other unions to organize their own feedback sessions so they could provide input.</p> <ul style="list-style-type: none"> • In the most recent teachers' contract, DCPS and the Washington Teachers Union (WTU) agreed to a new performance-based compensation system, called <u>IMPACTplus</u> <ul style="list-style-type: none"> - IMPACTplus allows teachers in the WTU who are rated "highly effective" to earn the maximum salary more quickly—nine years compared with 21 years in the old system—and increases the maximum salary a teacher can earn from \$87,500 to \$131,500 - Annual bonuses depend on teacher performance, whether the teacher teaches a tested grade/subject, the percentage of students eligible for free and reduced-price lunch (students in poverty) at the teacher's school, and whether the teacher teaches a high-need subject - Base salary increases are differentiated based on whether a teacher teaches at a high-poverty school - If teachers choose to accept the bonuses or salary increase, they must forgo some job security, including some options if a teacher is laid off and cannot find a position in another school - In 2010–11, approximately 70% percent of teachers eligible for bonuses accepted, and about 80% of teachers eligible for a salary increase accepted
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<p>DCPS plans to continue to review and improve the system over time, as appropriate.</p>
<p>GLOSSARY OF TERMS</p>	
<p><i>IMPACTplus</i></p>	<ul style="list-style-type: none"> • Performance-based compensation system that allows union members to earn a higher maximum salary, and to reach the maximum salary more quickly compared with the old salary schedule • Teachers forgo some job security if they accept the bonus or pay increase
<p><i>Master educator (ME)</i></p>	<ul style="list-style-type: none"> • Expert teaching practitioners whose full-time job is to travel from school to school as impartial evaluators • MEs provide teachers with support and feedback through written reports and post-observation conferences following each observation • MEs also support teachers in various other ways by participating in key district initiatives • MEs are content specialists, evaluating teachers who work in their particular content area whenever possible • MEs conduct about 200 observations a year and earn \$90,000+

TIMELINE



SOURCES

DCPS. *IMPACT—The DCPS effectiveness assessment system for school-based personnel*. Retrieved from <http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+%28Performance+Assessment%29>

DCPS. *IMPACT guidebooks*. Retrieved from <http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+%28Performance+Assessment%29/IMPACT+Guidebooks>

Headden, S. (2011). *Inside IMPACT: D.C.'s model teacher evaluation system*. Washington, D.C.: Education Sector. Retrieved from <http://www.educationsector.org/publications/inside-impact-dcs-model-teacher-evaluation-system>

ACHIEVEMENT FIRST (AF) TEACHER CAREER PATHWAY

OVERVIEW

Achievement First (AF) is a charter network that operated 20 schools in New York and Connecticut in the 2011–12 school year. In response to teacher feedback in 2009, AF committed to creating a Teacher Career Pathway (TCP) aimed at developing and celebrating great teachers by evaluating them against clear standards for instructional practice, and providing thoughtful and data-rich feedback to help teachers grow and develop their craft. After a year-long planning process involving teacher input groups and a five-month pilot, AF launched TCP in 2010–11.

TCP is based on Achievement First’s Teaching Excellence Framework. The framework consists of four components, broken into two “Inputs”—quality instruction and core values and contributions—and two “Outcomes”—student achievement and student character development. AF evaluates its teachers on these inputs and outcomes using student growth data, lesson observations, and peer, parent, student, and principal/dean surveys.

AF teachers are at-will employees and do not belong to a union. Therefore, the central focus of the system is to provide a framework for coaching and professional development for all AF teachers, and to identify and reward top performers. One of the main goals of the Teacher Career Pathway is to provide incentives for talented teachers to stay in the classroom. Teachers advance through five career stages, each associated with increased benefits, including compensation significantly above current traditional-district salary scales, network-wide public recognition, and select professional development opportunities. Advancing to new stages within the Teacher Career Pathway does not necessitate that teachers take on any additional responsibility.

AF recognizes that replication and scalability of TCP throughout the network is key to successful implementation. Although the Teacher Incentive Fund (TIF) currently covers costs for increased salaries of top teachers and schoolwide bonuses, AF is seeking to sustain the evaluation system within budget over time.

ABOUT THE SCHOOL SYSTEM

<i>Number of schools</i>	In 2011–12, 20 schools in two states serving students in grades K–12 <ul style="list-style-type: none"> • Nine schools in CT • 11 schools in NY
<i>Number of students</i>	6,200
<i>Number of teachers</i>	550

ABOUT THE SCHOOL SYSTEM

<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	Yes, from March–May 2009 with 33 teachers, and from January–May 2010 across all teachers in all schools
<i>In effect since</i>	2010–11
<i>Are there plans for additional phases/components?</i>	No

<i>Who gets evaluated?</i>	All teachers
<i>How often do teachers receive a rating?</i>	Once a year
<i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i>	Not applicable
STUDENT ACHIEVEMENT MEASURES	
<i>Assessments to measure student achievement</i>	<ul style="list-style-type: none"> • <u>New York State Test</u> - Grades 3–8 ELA and math • <u>New York High School Regents Examinations</u> - ELA, algebra I, geometry, algebra II, chemistry, physics, earth science, global history, U.S. history • <u>Connecticut Mastery Test</u> - Grades 3–8 reading and writing, math - Grades 5 and 8 science • <u>Connecticut Academic Performance Test (CAPT)</u> - Grade 10 reading, writing, math • <u>Fountas & Pinnell Benchmark Assessment</u> - K–2 ELA • <u>Terra Nova Assessment</u> - K–2 math • <u>SAT</u> and <u>PSAT</u> • <u>AP Exams</u> - U.S. history • <u>MCAS</u> - Physics • Network-wide Interim Assessments (IA) - Grades 6, 7, and 8 science IA - Grades 5–8 history IA - Grades 9–12 literature, composition, algebra I, geometry, algebra II, pre-calculus, global history, U.S. history I
<i>Are any new assessments in development to use with teacher evaluations?</i>	AF is currently working to improve the network’s history and science assessments. AF is also identifying a Spanish assessment and a new K–2 reading assessment.

SITE PROFILES

<p><i>What academic/achievement criteria are included?</i></p>	<p>AF uses one of three methods of measuring student achievement depending on the course: a teacher value-added model, a matrix model, or a portfolio review.</p> <ul style="list-style-type: none"> • Teacher-value added: For subjects and grades for which there are state standardized assessments, student growth is measured using a teacher value-added model that AF developed with Mathematica Policy Research. The model calculates typical growth based on results from both AF teachers’ classrooms and teachers in districts serving a comparable student population (New York City and Hartford). The student’s actual growth is compared to the typical growth of the comparable student group to determine the teacher’s value-added beyond typical growth • “Matrix model”: For subjects and grades for which a state standardized assessment does not exist, AF uses a “matrix model” that uses network-wide assessments to measure student growth from the prior year to the end of the year in comparison to typical growth for students at the same starting point • Portfolio process: For special classes (art, music, PE, dance, theater), AF uses a portfolio process: <ul style="list-style-type: none"> - In the fall, teachers work with a coach to decide on their program goals and assessments, including a written component and a skills-based assessment - In the spring, a content expert reviews progress towards student learning goals • The school principal can review student data and use other evidence of teacher performance to add points within a specific range to the “student achievement outcome” component <ul style="list-style-type: none"> - AF decided to grant principals more discretion because their close contact with coaches and teachers makes them very aware of whether the data best represent a teacher’s performance. In those instances, a regional superintendent will review the principal’s decisions - In courses where there are more informative student achievement measures, a principal has less discretion. In courses where there is a less-informative student achievement measure, a principal has more discretion
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<ul style="list-style-type: none"> • For teachers of state-tested grades/subjects, student achievement accounts for 40% of the rating • For teachers of untested grades/subjects, student achievement accounts for 20% of the rating
<p>NONACADEMIC MEASURES</p>	
<p><i>What nonacademic criteria are included?</i></p>	<ul style="list-style-type: none"> • The first nonacademic portion of the evaluation is called “Quality Instruction Input” and uses: <ul style="list-style-type: none"> - Lesson observations based on the AF Essentials of Great Instruction, AF’s set of 24 “elements” that outline instructional expectations for every teacher in the network • The second nonacademic portion of the evaluation is called “Student Character Development Outcome” and uses: <ul style="list-style-type: none"> - Student survey on their experience in the classroom (see p. 14 of Guidebook) - Family survey about character development and teacher’s relationships with students and families (See p. 14 of Guidebook) • The third nonacademic portion of the evaluation is called “Core Values and Contributions Input” and uses: <ul style="list-style-type: none"> - Teacher peer survey on core values and contributions to the mission (see p. 14 of Guidebook) - Principal survey of core values and contributions to the mission

<p><i>How much do the nonacademic measures count in a teacher's final rating?</i></p>	<ul style="list-style-type: none"> • For teachers of tested grades/subjects: <ul style="list-style-type: none"> - Student character accounts for 15% of the rating - Core Values accounts for 15% of the rating - Quality Instruction and Planning accounts for 30% of the rating • For teachers of untested grades/subjects: <ul style="list-style-type: none"> - Student character accounts for 15% of the rating - Core Values accounts for 15% of the rating - Quality Instruction and Planning accounts for 50% of the rating
<p>CLASSROOM OBSERVATIONS</p>	
<p><i>Who are the observers who rate teachers' effectiveness?</i></p>	<p>AF uses three types of observers:</p> <ul style="list-style-type: none"> • Coaches, who work closely with teachers throughout the year to help improve pedagogy • School leaders (principals and deans) • Instructional experts who are external to the school (AF regional superintendents, AF achievement directors, and some external consultants)
<p><i>What training do observers receive?</i></p>	<p>All observers attend training sessions at least once per quarter to view, rate, and discuss video lessons with other raters from across the network to norm their observations, debrief practices, and increase reliability.</p>
<p><i>Are observers compensated for this work?</i></p>	<p>No, observations are part of the observers' job description/salary.</p>
<p><i>How often are teachers observed?</i></p>	<ul style="list-style-type: none"> • Formal observations occur three times a year and last 45 minutes each • Teachers also receive frequent informal observations and feedback (approximately every other week). At the end of the year, evidence from informal observations are gathered into a "comprehensive lesson observation," a holistic view of the teacher's instruction
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<ul style="list-style-type: none"> • Informal observations: coaches observe teachers frequently throughout the year; no time frame • Teachers have one formal observation in the fall, one in the winter, and one in the spring
<p><i>Are observations announced or unannounced?</i></p>	<ul style="list-style-type: none"> • Informal observations: unannounced • Formal observations: unannounced
<p><i>What is the basis for the observation rubric?</i></p>	<ul style="list-style-type: none"> • The AF Essentials Rubric is based on the AF <i>Essentials of Effective Instruction</i>, which is informed by the work of Doug Lemov, Jon Saphier, and the observations of master teachers across the AF network • It was developed and tested during the 2010–11 school year and then revised and improved based on school leader and teacher feedback for the 2011-2012 school year

SITE PROFILES

<p><i>What is included in the observation scoring rubric?</i></p>	<p>AF's classroom <u>observation rubric</u> is based on the <u>Essentials of Effective Instruction</u></p> <ul style="list-style-type: none"> • Each "essential" is measured by performance indicators • Performance indicators are used to determine a rating for each essential based on a five-point scale: <ul style="list-style-type: none"> - Level 5 Exemplary: Consistently best-practice instruction that gives a high degree of confidence in breakthrough achievement gains - Level 4 Strong: Instruction aligned to best practices that gives strong confidence of achievement gains to consistently meet ambitious AF targets - Level 3 Solid: Solid instruction aligned to best practices that will likely lead to solid scholar achievement gains - Level 2 Emergent: Instruction that is mixed in quality of execution and may lead to scholar achievement concerns - Level 1 Ineffective: Instruction that could lead to very serious scholar achievement concerns • Each essential is then weighted into an overall observation score between 10 and 100
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • Informal observations: Principal keeps coaches accountable through biweekly meetings and reports of how coaches/teachers are doing • Formal observations: In addition to the required training outlined above, observers are also held accountable by their co-observers, by the teachers they observe, and by the network <ul style="list-style-type: none"> - Observations include two observers co-observing a lesson and discussing the evidence they have to support a rating - Teachers complete an optional feedback survey after each observation, which includes feedback on their observers - At the network level, AF watches for trends in feedback as well as trends in the actual ratings observers assign to identify outliers
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • Informal observations: Coaches debrief every observation with their teachers. Feedback is usually oral, but may also be written • Formal observations: Each observation is debriefed within one week of the observation through written feedback and an in-person conversation. This formal conversation is in person and generally includes all observers who saw the lesson. Teachers provide feedback on the debrief conversation through an optional post-observation feedback survey

ACCURACY, VALIDITY, AND RELIABILITY	
<i>Is there a process to validate non-standardized assessments?</i>	Yes, all assessments are network-wide except for “specials” teachers (art, music, PE, dance, theater); these teachers’ assessments are reviewed in the fall by the network’s Specials Achievement Director.
<i>Are there procedures to ensure that observation scores are valid?</i>	See “How are observers held accountable for their scoring?” above.
<i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i>	<ul style="list-style-type: none"> • When the value-added results align clearly with other results (class observations, survey results), a teacher’s final evaluation score is finalized • When there is a misalignment between the value-added results and other results, the school’s regional superintendent will review a teacher’s student achievement outcomes, then review the teacher’s placement into a stage. This happens rarely, as the system is designed to incorporate multiple parties’ input into a teacher’s rating
<i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i>	See above
USING AND REPORTING EVALUATION RESULTS	
<i>What are the rating categories?</i>	<p>Teachers receive a rating on a 1 to 5 scale. Each rating correlates with a different stage in AF’s Teacher Career Pathway.</p> <ul style="list-style-type: none"> • Stage 1—Intern: An intern works with small groups, helps to check homework, monitors lunch and recess, etc. An intern is not responsible for his or her own class. Interns can advance to Stage 2 after one successful year, at the discretion of the principal • Stage 2—Teacher: Solid contributor, rapidly developing; delivers solid student achievement. Stage 2 teachers can advance after two years of successful results, at principal’s discretion • Stage 3—Teacher: Strong, stable contributor; delivers strong student achievement. Stage 3 teachers can advance after two years of highly successful results • Stage 4—Senior Teacher: Strong, stable contributor; delivers very strong student achievement; meets rigorous requirements. Stage 4 teachers can advance after two years of exemplary results • Stage 5—Master Teacher: Exceptional contributor; consistently exemplary student achievement; meets rigorous requirements
<i>How are results reported?</i>	<ul style="list-style-type: none"> • Individual teacher results are reported to teachers, their coaches, and school leaders only • Individual teacher evaluation results will not be available publicly • Results are currently paper-based, but in 2012–13, AF plans to move to a talent-management portal that allows teachers and coaches to have access to all of the data included in a teacher’s final evaluation as it is available

SITE PROFILES

<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<p>Yes, evaluation results are tied to professional development, pay, and recognition.</p> <ul style="list-style-type: none"> • Professional development <ul style="list-style-type: none"> - Teachers at stages 4 and 5 become part of senior teacher cohort and receive a self-directed professional development budget, opportunities to visit excellent teachers nationally, and preferred access to special PD sessions - For low-performing teachers, coaches may place teachers on a performance improvement plan (PIP) at any time, which outlines specific requirements/actions the teacher must complete within 4–8 weeks to renew contract. Teachers are continually supported by coaches while on the improvement plan • Pay <ul style="list-style-type: none"> - Teachers are eligible to move up the career pathway stages described above under “rating categories.” Moving up the pathway does not require a teacher to take on more responsibility (although such opportunities exist), but it does increase pay. - Two consecutive years of poor performance can cause a teacher to move to a lower stage - AF’s system aims to encourage professional growth while also providing an incentive for great teachers to stay in the classroom and keep teaching <ul style="list-style-type: none"> » Teachers at stages 4 and 5 will continue to receive annual pay raises » Teachers at stages 2 and 3 will see annual pay raises for a limited time, but their salaries will eventually flatline if they do not progress to the next stage • Increased responsibility <ul style="list-style-type: none"> - At their discretion, principals may offer teachers additional leadership opportunities (e.g., grade-level team leader, coach, dean) to excellent teachers with a stipend • Other (for stage 4 and 5 teachers only) <ul style="list-style-type: none"> - Recognition of those moving up through stages through school-based “appreciations” - Announcements in network-wide newsletter - Videos of teaching used as exemplars - Special luncheons with network leaders
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<ul style="list-style-type: none"> • A teacher may submit a “request for review” to the review committee (principal, regional superintendent, member of AF Network Support Team) • Reviews can be requested regarding three aspects of the Teacher Career Pathway. Teachers can request a review of their eligibility status to participate in the Teacher Career Pathway; of an observation that occurred that they thought didn’t meet the standards for an “observable lesson”; or of a stage advancement decision • Committee will make final decision
<p><i>If performance results are available, what is the score distribution?</i></p>	<p>Results for the overall teacher excellence framework evaluation are not yet available. For the first observation window in fall 2011, on a 10- to 100-point scale that scores the observation based on the <i>Essentials of Effective Instruction</i>, the average rating was a 52.</p>

TEACHER INVOLVEMENT AND CONTINUING IMPROVEMENT

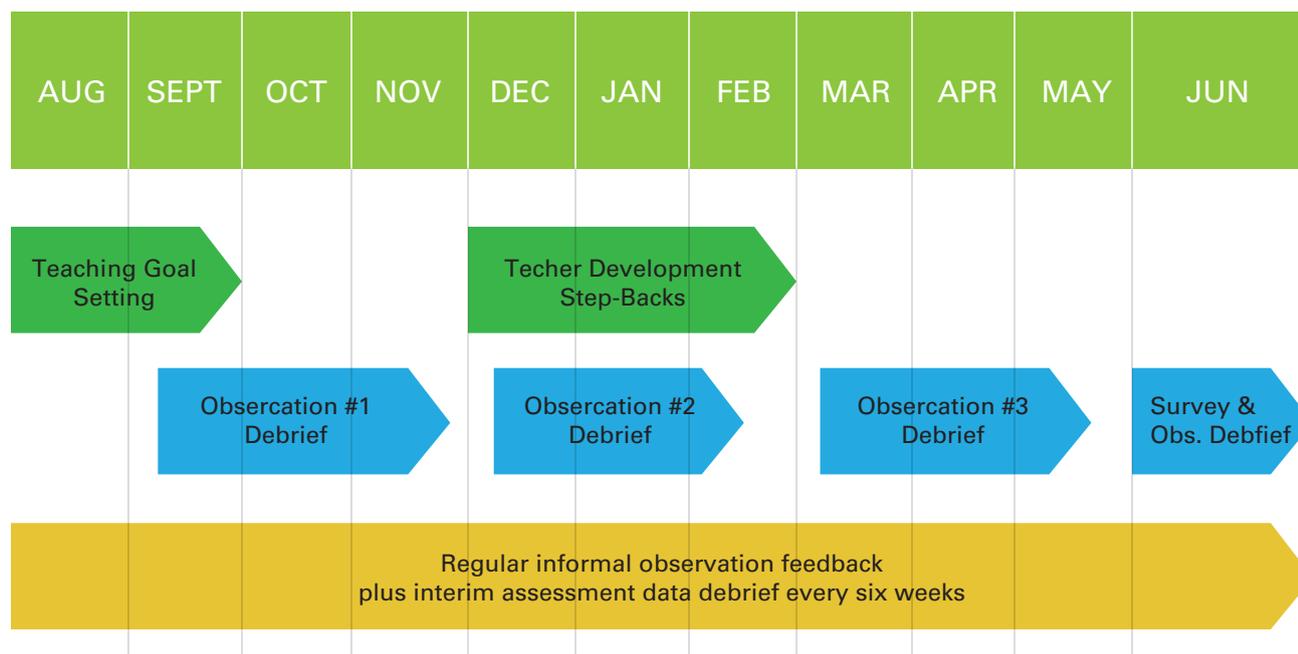
<p><i>What role did teachers play in developing the evaluation system?</i></p>	<ul style="list-style-type: none"> • In 2008, teachers, academic deans, and principals worked for one year to develop the AF <u>Essentials of Great Instruction</u> • In 2009–10, teachers and AF network leaders developed the components of the Teacher Career Pathway together. Discussions included: <ul style="list-style-type: none"> - What the four components should be - How to measure components - Survey questions for parents, students, peers, and school leaders - Rewards for teachers - Stages on the career pathway • In 2010–11, teacher-input groups for each “specials” subject area developed a unique portfolio review process to assess student achievement outcomes in specials courses • Teachers have continued input every year: All teachers provide feedback via multiple surveys, and two representative teachers from each school (nominated by principals) meet every six weeks to give input on the Teacher Career Pathway
<p><i>What role did the union play in developing the evaluation system?</i></p>	<p>Not applicable</p>
<p><i>Is there a formal process for revising the evaluation system over time?</i></p>	<p>Revisions are constantly made to the Teacher Career Pathway using feedback from representatives at each school (teacher advisory panel), post-lesson observation feedback surveys, school leader feedback, operations team feedback, and other annual surveys</p>

GLOSSARY

<p><i>Coach</i></p>	<ul style="list-style-type: none"> • Can be principal, academic dean, or expert teacher. Coach works closely with teachers throughout the year to help improve pedagogy
<p><u>Essentials of Great Instruction</u></p>	<ul style="list-style-type: none"> • Set of 24 “elements” that outline instructional expectations for every teacher in the network

SITE PROFILES

TIMELINE



SOURCES

Achievement First. *AF teacher career pathway executive summary*.

Achievement First. (2011). *Teacher guidebook 2011–2012: Achievement First Teacher Career Pathway*.

Achievement First. (2011). *AF Teacher Career Pathway: 2011–2012 student achievement measures*.

Achievement First. (2011). *AF Essentials observation rubric: 2011–2012*.

Curtis, R. (2011). *Achievement First: Developing a teacher performance management system that recognizes excellence*. Washington, DC: The Aspen Institute. Retrieved from http://www.aspeninstitute.org/sites/default/files/content/docs/education%20and%20society%20program/AI_Achievement%20First_performance%20mangmt.pdf

RELAY GRADUATE SCHOOL OF EDUCATION

OVERVIEW

In 2008, the leaders of three leading charter management organizations—Achievement First, Knowledge is Power Program (KIPP), and Uncommon Schools—partnered with the Hunter College School of Education in New York City to create Teacher U, a two-year master’s program for novice teachers, grounded in proven, practical techniques that teachers can immediately use in their classrooms to lead their students to higher achievement. In February 2011, the New York State Board of Regents granted Teacher U’s leaders a charter to operate an independent institution of higher education. Relay Graduate School of Education (GSE) opened its doors in July 2011 as the first stand-alone school of education to be newly credentialed in New York in more than 80 years.³⁵

Today, approximately 420 novice teachers are enrolled in Relay GSE’s Master of Arts in Teaching (MAT) program, honing their craft as they work as full-time K–12 teachers. To be considered for admission, applicants must have secured a full-time classroom teaching position for the upcoming academic year and must not be certified to teach in New York State. Due to current hiring conditions in New York City, most of the teachers who attend Relay GSE work in public charter schools.

Although Relay GSE is not a K–12 school system or CMO, we include it in our review because its graduation requirements function much like an evaluation system. Graduate students in the program are observed two to three times a year for two years, assessed on their mastery of 60 pragmatic modules related to classroom practice, and must demonstrate at least a year’s worth of K–12 student growth to graduate from the program. Teachers submit their student achievement results through the Master’s Defense, a portfolio created during the second year of the program that showcases their proficiency across five elements of effective teaching.

Though Relay GSE was just established in 2011, two cohorts of teachers have completed the Teacher U at Hunter College program and submitted Master’s Defenses. The third and final Teacher U at Hunter College cohort will complete the program in the summer of 2012. Relay GSE’s first cohort of graduate students will graduate in the summer of 2013.

Standout features of Relay GSE’s graduation requirements include a pre-approved list of assessments by which to measure student growth across a range of grades and subjects, and a random review of teachers’ documentation of student performance each year. Interviewees identified the program’s greatest challenges as scaling up the accountability systems, which can take much time for professors, developing a valid and rigorous assessment item bank for all grades and subjects, and aligning its tools and trackers with those used in its teachers’ schools.

ABOUT THE SCHOOL SYSTEM³⁶

Number of schools	93
Number of students	21,000 (estimated)
Number of teachers	420

³⁵ Please note: The program described below is that of Relay Graduate School of Education. The Relay philosophy, curriculum, and approach are very similar to that of Teacher U at Hunter College (TUHC). Although the policies and procedures described below are currently in place at Relay GSE, the graduate student and program performance data comes largely from the TUHC program, because Relay GSE is still in its first year of operation.

³⁶ Relay GSE is not a traditional school system. The statistics describe the teachers enrolled in the program, the number of schools in which they teach, and the number of students that they teach.

SITE PROFILES

BACKGROUND	
<i>Corresponding legislation</i>	Not applicable
<i>Year legislation passed</i>	Not applicable
<i>Was the system piloted?</i>	Since TUHC's first graduating class in 2010, teachers have had to demonstrate measurable student growth to receive a degree, although the requirements to do so have evolved.
<i>In effect since</i>	Teacher U (developed in concert with Hunter College) became Relay Graduate School of Education (Relay GSE) in 2011–12.
<i>Are there plans for additional phases/components?</i>	Beginning in 2012-13, second-year graduate students must demonstrate student growth in two subjects/classes (if more than one is taught).
<i>Who gets evaluated?</i>	All teachers enrolled at Relay GSE
<i>How often do teachers receive a rating?</i>	<ul style="list-style-type: none"> • At the end of Relay GSE's two-year MAT program, teachers receive a single score out of 100 graduation points which come from points awarded across four categories: Student Growth & Achievement, In-Person Observations, Program Success, and Master's Defense Portfolio • Teachers must receive a score of 70 or better to graduate • Teachers receive formative feedback on points earned to date throughout the two years enrolled in the program
<i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i>	Not applicable

STUDENT ACHIEVEMENT MEASURES

<p><i>Assessments to measure student achievement</i></p>	<p>Relay GSE has created a list of approved measures of student achievement for all grades and subjects:</p> <ul style="list-style-type: none"> • Reading <ul style="list-style-type: none"> - <u>Developmental Reading Assessment 2</u> (grades K–8) - <u>Fountas and Pinnell Benchmark Assessment System</u> (grades K–8) - <u>Reading A to Z</u> (grades K–8) - <u>Rigby</u> (grades K–3) - <u>STEP Reading Assessment</u> (grades K–3) - <u>Teachers College Reading & Writing Project Assessment</u> (grades K–8) • English language arts <ul style="list-style-type: none"> - Relay GSE-approved, teacher-constructed, or teacher-acquired ELA assessments showing mastery of state standards - Relay GSE-approved, teacher-constructed, or teacher-acquired writing rubric(s) showing mastery/growth on state standards or approved writing traits • Math <ul style="list-style-type: none"> - Relay GSE approved, teacher-constructed, or teacher-acquired math assessments showing mastery of state standards • Science <ul style="list-style-type: none"> - Relay GSE-approved, teacher-constructed, or teacher-acquired science assessments showing mastery of state standards • Social Studies <ul style="list-style-type: none"> - Relay GSE-approved, teacher-constructed, or teacher-acquired social studies assessments showing mastery of state standards • Other subjects <ul style="list-style-type: none"> - Relay GSE-approved, teacher-constructed, or teacher-acquired assessments and/or rubrics showing mastery/growth on state or national standards <p>Note: Relay GSE would be interested in using state standardized assessments to measure student growth. However, because few teachers (25 percent) teach classes with standardized tests, and the results aren't available until the following summer or autumn, Relay GSE does not currently use state standardized assessments.</p>
<p><i>Are any new assessments in development to use with teacher evaluations?</i></p>	<ul style="list-style-type: none"> • Relay GSE is hoping to build an assessment item bank • Rather than creating new assessments, Relay GSE aims to include more rigorous, externally validated assessments that already exist • The vetting process for a new assessment includes: <ul style="list-style-type: none"> - Faculty review for rigor and alignment - If using multiple interim assessments: should assess each academic standard at least three times - If using a single summative assessment: should assess at least half of academic standards (with at least two questions addressing each of the standards) - Exceptions made on a case-by-case basis

SITE PROFILES

<p><i>What academic/achievement criteria are included?</i></p>	<p>In their second year, all teachers are held accountable for meeting a minimum floor for student achievement in their classroom in two subjects or classes (e.g., different sections or grades of the same subject, where applicable) as a way of holding teachers accountable for student gains in the majority of the core curriculum. At the same time, teachers can receive a higher score by exceeding the floor and pursuing an “ambitious goal” for student growth:</p> <ul style="list-style-type: none"> • The “achievement floor” is at least a year’s worth of growth or 70% standards mastery <ul style="list-style-type: none"> - Assuming students are behind grade level, making one year of growth won’t close the achievement gap. Ultimately, Relay GSE would like to hold their teachers accountable for more growth, but now uses a single year as a floor to accommodate the inevitable variation associated with a single year of achievement data - If teachers don’t meet the floor in at least one subject, they get zero points for this component and will not earn a degree • The “ambitious goal” is based on students’ diagnostic scores for elementary reading and 80% mastery for other subjects • See p. 8 in Graduation Requirements, Class of 2013 for more detail
<p><i>How much do the student achievement measures count in a teacher’s final rating?</i></p>	<p>Student achievement counts for 45 percent of a teacher’s final score.</p> <ul style="list-style-type: none"> • In Year 1, teachers earn up to 5 points (out of 100) if they know the level of their students’ academic achievement • Student growth in Year 2 counts for as many as 40 graduation points out of 100. Teachers get credit if they meet the floor or ambitious goal in one subject, but can earn more points if they meet it in more subjects. For example: <ul style="list-style-type: none"> - If teachers don’t at least meet the floor in one subject, they earn no graduation points for this component - If teachers meet the floor in one subject or grade level, but not the second, they earn 25 points - If teachers meet the floor in both subjects, they receive 30 points - Teachers receive an additional 5 points for each subject in which they reach the ambitious goal - See pp. 6–8 in Graduation Requirements, Class of 2013 for more detail
<p>NONACADEMIC MEASURES</p>	
<p><i>What nonacademic criteria are included?</i></p>	<p>Relay GSE considers four nonacademic measures:</p> <ul style="list-style-type: none"> • Program success (includes passing 60 course modules, and professionalism) • Classroom observations • Master’s Defense, a portfolio created during the second year of the program that contains student growth and achievement information and showcases proficiency across five elements of effective teaching • Student character growth in Year 2 <ul style="list-style-type: none"> - Includes qualities such as grit, self-control, gratitude, zest, and humor - Teachers will assess students’ character using a rubric, and receive points toward graduation accordingly - For more details on holding teachers accountable for character, see “What if the secret to success is failure?”

<p><i>How much do the nonacademic measures count in a teacher's final rating?</i></p>	<p>Nonacademic measures count for 55 percent of a teacher's final score:</p> <ul style="list-style-type: none"> • Program success— 30 points • Classroom observations—12 points • Master's defense—8 points • Character growth—5 points
<p><i>Can LEAs devise their own effectiveness measures where not already defined in legislation?</i></p>	<p>Not applicable</p>
<p>CLASSROOM OBSERVATIONS</p>	
<p><i>Who are the observers who rate teachers' effectiveness?</i></p>	<p>Relay GSE faculty. All faculty members are recent K–12 teachers with track records of closing the achievement gap in their classrooms.</p>
<p><i>What training do observers receive?</i></p>	<p>Faculty members observe graduate students using a common rubric, and all faculty members are aligned on observation protocols after receiving 1 to 2 days of norming.</p>
<p><i>Are observers compensated for this work?</i></p>	<p>Observations are part of faculty member's job description/salary.</p>
<p><i>How often are teachers observed?</i></p>	<ul style="list-style-type: none"> • Five formal observations over two years <ul style="list-style-type: none"> - Three in Year 1 - Two in Year 2 - The first observation does not count towards graduation
<p><i>Is there a time frame within which observations must be conducted?</i></p>	<ul style="list-style-type: none"> • Year 1 <ul style="list-style-type: none"> - August–September - Fall - Spring • Year 2 <ul style="list-style-type: none"> - Fall - Spring
<p><i>Are observations announced or unannounced?</i></p>	<p>Announced</p>
<p><i>What is the basis for the observation rubric?</i></p>	<p>Relay GSE has developed its own observation rubrics, but they are based on the following work/rubrics: Teach for America's <i>Teaching as Leadership</i>, Charlotte Danielson's <i>Framework for Teaching</i>, Kim Marshall's <i>Rethinking Teacher Supervision and Evaluation</i>, Robert Piant's <i>Classroom Assessment Scoring System (CLASS)</i>, and Doug Lemov's <i>Teach Like a Champion</i>.</p>
<p><i>What is included in the scoring rubric?</i></p>	<ul style="list-style-type: none"> • Observers look for six things: <ul style="list-style-type: none"> - Student engagement - Use of time toward objective - Classroom management - Lesson alignment - Checking for understanding - Questioning <p>See pp.16–18 in Graduation Requirements, Class of 2013 for more detail.</p>

SITE PROFILES

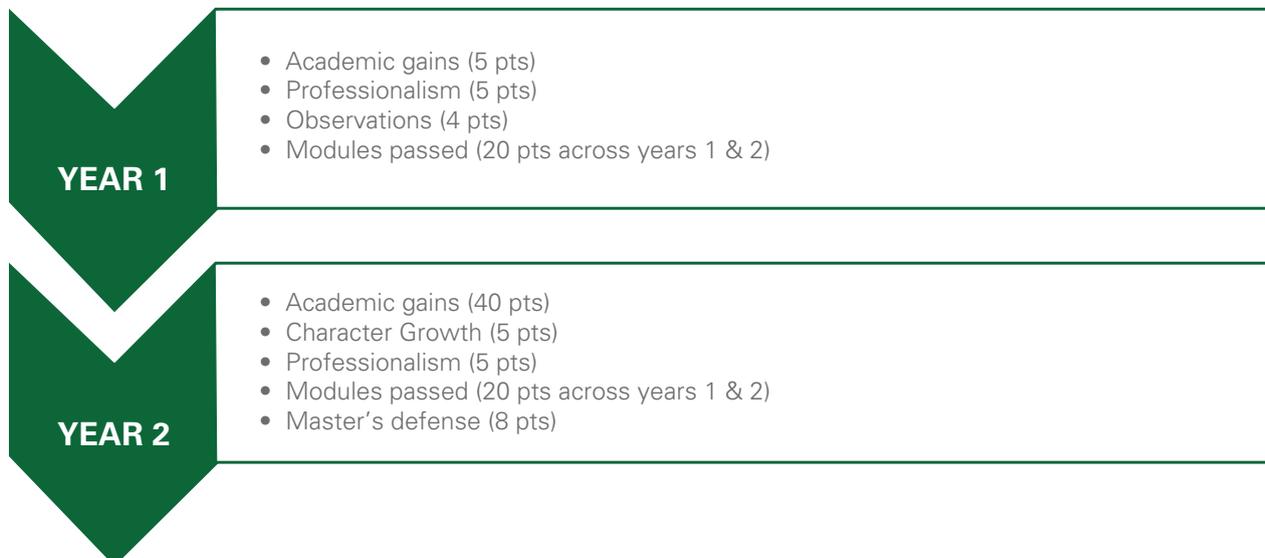
<p><i>How are observers held accountable for their scoring?</i></p>	<ul style="list-style-type: none"> • Observers receive 1 to 2 days of norming to ensure that observers' scoring is consistent, and practice with the rubric/observation tools • Each observation is logged in a centralized data system, and Relay's director of research analyzes differences in observational data that are likely artifacts of the observer and not the teacher • Such differences result in exploratory conversations between the observer and the dean of teaching and learning about what might be associated with these differences (e.g., lack of knowledge of the rubric, school context, etc.)
<p><i>Do teachers have an opportunity to debrief with observers?</i></p>	<ul style="list-style-type: none"> • Yes, students at Relay GSE receive rubric-based feedback from the assistant professor observing them in writing, followed by a half-hour in-person or phone conversation focused on the rubric criteria • The debrief session provides graduate students with an opportunity to both ask questions about the feedback and determine next steps
<p>ACCURACY, VALIDITY, AND RELIABILITY</p>	
<p><i>Is there a process to validate non-standardized assessments?</i></p>	<ul style="list-style-type: none"> • Relay GSE sets criteria for all assessments used to measure student growth, and faculty must approve the assessments • Every year, Relay GSE conducts a random review of 5 to 10 percent of second-year students <ul style="list-style-type: none"> - During the review, Relay GSE reviews all documentation and has a conference with the teacher to learn about his/her students' work - Examples of documentation may include: <ul style="list-style-type: none"> » Assessment map » Running record sheets » Student work samples - Thus far, 100 percent of Relay GSE teachers have passed the random review • An external party (Achievement Network) has been contracted to evaluate teacher-designed assessments used to make inferences about standards mastery • Relay GSE also monitors teachers' performance on an ongoing basis <ul style="list-style-type: none"> - Relay GSE faculty members are responsible for overseeing groups of 25 to 50 teachers, monitoring their progress, and providing support - Small groups allow faculty to know teachers and their classrooms well • If performance or documentation seems insufficiently aligned, faculty will follow up as needed
<p><i>Are there procedures to ensure that observation scores are differentiated?</i></p>	<p>No</p>
<p><i>Is there a procedure to validate that observation scores are correlated with student outcomes?</i></p>	<p>At the end of the 2011–12 school year, Relay GSE's director of research will conduct the first analyses correlating observation scores with student outcomes.</p>

<p><i>Is there a procedure to validate that other nonacademic measures are correlated with student outcomes?</i></p>	<p>At the end of the 2011–12 school year, Relay GSE’s director of research will conduct the first analyses correlating student outcomes with other Relay data, including (but not limited to):</p> <ul style="list-style-type: none"> • performance on module-specific rubrics • professionalism scores • attendance at class sessions • undergraduate GPA
<p>USING AND REPORTING EVALUATION RESULTS</p>	
<p><i>What are the rating categories?</i></p>	<p>Teachers earn up to 100 points, and must earn 70 points or more to receive a master’s degree.</p>
<p><i>How are results reported?</i></p>	<p>To teacher only</p>
<p><i>Are there consequences tied to evaluation results? If yes, what are they?</i></p>	<p>Graduation and Relay GSE’s recommendation for certification are tied to the evaluation results.</p> <ul style="list-style-type: none"> • Teachers must earn at least 70 out of 100 points to graduate and earn a master’s degree from Relay GSE • It is impossible to graduate without meeting the achievement floor in at least one subject • In addition, Relay GSE bases its recommendations for teacher certification on successful completion of its graduate program
<p><i>Can teachers appeal their ratings? If yes, how?</i></p>	<ul style="list-style-type: none"> • Yes, but no appeals took place during the first two cohorts of graduate students at Teacher U at Hunter College (the predecessor to Relay GSE) • If teachers do appeal their score, they will follow the random review process
<p><i>If performance results are available, what is their score distribution?</i></p>	<ul style="list-style-type: none"> • On average, second-year teachers at Teacher U at Hunter College lead their students to make 1.3 years of growth in reading: <ul style="list-style-type: none"> - 40% of teachers made 1.5 years of growth - 60% of teachers made between 1.0 and 1.5 years of growth
<p><i>If results are used to place a teacher on an improvement plan or dismiss a teacher, what does that process involve?</i></p>	<ul style="list-style-type: none"> • Graduate students are in regular contact with their assistant professor to discuss progress in the program. Should a graduate student need academic assistance, the student’s assistant professor provides academic support, additional feedback and/or resources to help the student improve his/her performance in the program • In addition, any graduate student whose pace of academic progress is unsatisfactory despite the additional supports provided by his/her assistant professor will be placed on academic probation <ul style="list-style-type: none"> - Graduate students who are placed on probation will be required to have an initial meeting with their faculty members to develop a written plan to return to good academic standing. The plan must be signed by the graduate student and the faculty member and filed in the Office of Enrollment Services. The plan must address how the graduate student will successfully complete the program in three years - Graduate students have one term to return to satisfactory academic standing; failure to do so will result in dismissal from Relay GSE - The dean and the faculty will review the academic performance of the graduate student before deciding whether to dismiss a graduate student. Graduate students who have been dismissed can petition their dismissal by contacting the dean to request reinstatement. The dean will make a decision regarding reinstatement. The dean’s decision is final

SITE PROFILES

TEACHER INVOLVEMENT	
<i>What role did teachers play in developing the evaluation system?</i>	Graduate students were not directly involved in developing the graduation system. However, Relay GSE solicited input from teachers and school leaders at Achievement First, KIPP, and Uncommon Schools.
<i>What role did the union play in developing the evaluation system?</i>	Not applicable
<i>Is there a formal process for revising the evaluation system over time?</i>	Relay GSE continually refines its assessment process and implements major revisions during the summers between academic years.
GLOSSARY OF TERMS	
<i>Modules</i>	<ul style="list-style-type: none"> • Units of study pertaining to classroom practice • Students must demonstrate mastery of 60 modules
<i>Achievement floor</i>	<ul style="list-style-type: none"> • The minimum student growth required to meet Relay GSE's graduation requirements • Set at 1 year's growth or 70% standards mastery, on average
<i>Ambitious goals</i>	<ul style="list-style-type: none"> • Student achievement goals that are both ambitious and feasible based on student performance on diagnostic test or the previous year's test • Exceeds one year's growth • 80% standards mastery or better

TIMELINE



See *Graduation Requirements, Class of 2013* for more detail.

SOURCES

RELAY/GSE. (2011). Graduation requirements, Class of 2013.



**MORE
INFORMATION**

ABOUT THE AUTHORS

DANIELA DOYLE is a consultant with Public Impact. Her work covers a range of issues, including teacher quality, school finance, charter schools, turnarounds, and redesigning school models and teachers' roles to extend the reach of excellent teachers. A former elementary school teacher, Ms. Doyle is an alumna of Teach For America and Education Pioneers, a non-profit bringing innovative leaders to education. She holds a bachelor's degree in public policy from Princeton University, a master's of science for teachers from Pace University, and a master's degree in public policy from Duke University.

JIYE GRACE HAN is a policy fellow with Public Impact. Her work addresses a variety of topics, including retaining and extending the reach of high-performing teachers, teacher and leader quality, finance reform, and technology in education. Before joining Public Impact, Ms. Han served as a Teach For America corps member working in a high-poverty school. In 2011, she was named a national finalist for the Sue Lehmann Excellence in Teaching Award, given annually to second-year Teach For America corps members who embody and demonstrate the leadership needed to have an exemplary level of transformational impact with students. She has a B.S. in journalism from Northwestern University and a master's in teaching from Dominican University.

ABOUT 50 CAN

The 50-State Campaign for Achievement Now identifies and supports local leaders building movements within their states to make sure that every child has access to a great public school. We believe that great schools change everything. www.50can.org

ABOUT CONNCAN

ConnCAN – the Connecticut Coalition for Achievement Now – is leading a movement to improve educational outcomes for Connecticut's kids. We bring advocates, policy makers, parents, educators, and people like you together to change the system and give all kids access to great public schools.

Since 2005 we've been hard at work doing just that. And the opportunity has never been greater to fundamentally transform the way we deliver on the promise of a great public education.

www.conncan.org

ABOUT PUBLIC IMPACT

Public Impact's mission is to dramatically improve learning outcomes for all children in the U.S., with a special focus on students who are not served well. A national education policy and management consulting firm based in Chapel Hill, NC, Public Impact is a team of researchers, thought leaders, tool-builders, and on-the-ground consultants working with leading education reformers. For more on Public Impact, please visit: www.publicimpact.com.