



Making High-Quality Teacher Evaluation Manageable

July 2013

Managing a New Evaluation System: A Principal's View

For William "Jud" Haynie, a principal in Metro Nashville Public Schools (MNPS), implementation of the new Tennessee Educator Acceleration Model (TEAM) has not been the overwhelming experience he and others might have predicted. The school district's commitment to teacher evaluation as a tool for improving instruction lets Haynie focus his time and energy on his most important tasks — observing teachers and giving them quick, constructive feedback. Though the new system requires him to conduct more observations, he doesn't find this new expectation unmanageable.

MNPS used the move to a new evaluation system to streamline the work of principals. The district dropped many tasks demanded by the prior system, including time-consuming paperwork and district meetings that required significant time away from his school. In 2010–11, the pilot year for TEAM, MNPS made a clear commitment to supporting principals in their new roles. Mr. Haynie recalls that when principals said they needed more time in classrooms and help conducting observations, the district responded by releasing them from district-wide commitments and hiring retired administrators on part-time contracts to provide back-up administrative support in buildings, conduct observations and serve as instructional coaches.

—Discussion with Jud Haynie, May 12, 2012

Introduction

State education agencies (SEAs) and local educational agencies (LEAs) implementing evaluation reforms aspire to topnotch educator evaluation systems that are as manageable as Principal Haynie's is for him. They frequently report tension, however, between two values. On the one hand, they want to make sure the system is high quality — that it rigorously measures excellent practice, incorporates evidence of student learning and serves as a means to support the development of all practitioners. On the other hand, they want it to be manageable — something that can be done well without an exorbitant allocation of resources such as time, money and staffing.

"The Teacher Evaluation and Development Process endeavors to shift the national paradigm of thinking on teacher evaluation to one in which both administrators and teachers approach individual and team-based development as the highest priority to achieve student learning goals."

—*New Haven Public Schools Teacher Evaluation and Development Process: Teachers and Administrators Guide*

The Reform Support Network, sponsored by the U.S. Department of Education, supports the Race to the Top grantees as they implement reforms in education policy and practice, learn from each other, and build their capacity to sustain these reforms, while sharing these promising practices and lessons learned with other States attempting to implement similarly bold education reform initiatives.

It would be easy to resolve this tension by trading one value for another. To make the system more manageable, sacrifice quality. To raise the quality of the system, accept that it will be difficult to manage. With ingenuity and commitment to durable change, some States and LEAs, however, are finding they can avoid these tradeoffs. MNPS made Mr. Haynie's job more manageable by rethinking his duties, expanding the ranks of those who conduct observations, providing more administrative support in buildings and delivering additional instructional coaches to address teacher development needs uncovered in the evaluation process.

Metro Nashville's ability to make adjustments in response to educator feedback benefits from a statewide culture of continuous improvement fostered by the Tennessee State Department of Education. In fact, its staff meets regularly with LEA leaders, principals and teachers to gather information and solicit ideas on how to improve TEAM. This Volunteer State practice led to policy revisions reducing the number of observations required annually for effective teachers, a change discussed elsewhere in this publication.

The Role of SEAs and LEAs

That this publication begins with an explanation of how a State and one of its major urban school districts have focused on making evaluation manageable while maintaining quality is purposeful. Delineating the roles played by SEAs and LEAs is a challenge that leaders can confront early as they grasp the unique contribution each can make.

Evaluation that leads to improved teaching and learning ultimately depends on the relationship between an evaluator and the person being evaluated. LEAs are closer to this relationship, because both of these people usually are their employees. As a consequence, many of the successful approaches to making evaluation more manageable described in this publication come from LEAs, though there is no reason States cannot implement similar practices.

At the same time, States have their own unique role to play and can tailor supports to LEAs based on

their size, resources, and even their vision and will. Tennessee and other Race to the Top grantees often have policies that create model evaluation frameworks — including model rubrics for observations — which LEAs must use, unless they successfully argue for a waiver by proving their own locally developed systems are just as rigorous as the State's. With bigger budgets and readier access to philanthropic resources, large urban and suburban districts can pursue these waivers more readily, develop their own approaches and refine them over time. Smaller districts often cannot do the same. Denver Public Schools can break new ground, designing, piloting and implementing its own unique evaluation system, Leading Effective Academic Practice (LEAP), with support from local and national philanthropy. A small mountain district on Colorado's western slope, however, is going to have to rely on the Colorado Department of Education, some 200 miles away, for guidance and support.

SEAs can support these very different types of school districts using practices pioneered by leading States and school districts. They can also rethink how they function. Where large districts can purchase third-party assistance using philanthropic and other resources, smaller districts with fewer resources must rely on the State to function like these third-party entities. In this case, the State provides technical assistance for LEAs as individuals and groups. In addition, the SEA makes available high-quality resources for districts to employ themselves.

The remainder of this paper captures States and LEAs in the act of making their systems manageable and of high quality. It is divided into two parts. The first explores quickly deployable, high-impact tactics that these jurisdictions are using to make their evaluation systems more manageable. The second examines the foundation upon which States and school districts must build high-quality, manageable evaluation systems: a commitment to aligning college- and career-ready standards implementation with evaluation reforms and educator engagement.

Employing High-Impact Tactics

High-impact tactics allow leaders to make relatively quick adjustments to policy or practice. To maintain

quality but make their evaluation systems more manageable, SEAs and LEAs have employed tactics that include broadening the responsibility for observations, redefining school leadership, deploying feedback loops, balancing thoroughness with efficiency and using appropriate technology.

Sharing the Responsibility for Observation

States and school districts can increase the number of observers by expanding the categories of professional staff allowed to conduct evaluations. Historically, States and school districts have permitted only those with administrative licenses to observe teachers. Those administrators typically observed teachers infrequently, sometimes not at all. The observation load was manageable.

Now, however, many States and school districts have increased the required number of classroom observations so that they can provide adequate support to teachers and make more accurate judgments about their performance. Increasing the number of observers relieves the burden of additional observations. It has the added bonus of increasing the reliability of observations as a component of teacher evaluation systems.¹

Laws or rules in many States permit districts to use teachers as peer observers. In Hillsborough County Public Schools in Florida and Denver Public Schools in Colorado, teachers play this role in the evaluation process, with their observations counting toward their colleagues' overall rating. Ohio also allows districts to use peer evaluators. In 2012, the State trained teachers alongside administrators as it began the process of certifying the large pool of evaluators required to implement its new evaluation system.

New Haven Public Schools in Connecticut has its own approach to increasing the number of observers. District and union leaders worked together to craft a new evaluation system that recognizes two categories: *instructional managers*, who are principals, assistant

¹ Andrew Ho and Thomas Kane. "The Reliability of Classroom Observations by School Personnel" (The Measures of Effective Teaching [MET] Project, Bill & Melinda Gates Foundation, January 2013) 29.

principals or other administrators whom the teacher and the principal recognize as accountable for leading that teacher's evaluation and development, and *third-party validators*, who are retired teachers not affiliated with the schools in which they observe teachers. This approach allows principals to build evaluation teams, bring in additional observers to lighten their own observation load, and increase the validity of the conclusions instructional managers reach about teacher effectiveness.

As Katya Levitan-Reiner, New Haven's former senior coordinator in the Department of Research, Assessment and Student Information, explained, "The system gives school leaders flexibility, and it allows a different school leadership structure to emerge in different settings." In one school she described, the principal has been thoughtful about building a team of instructional managers. Instead of trying to observe every teacher, she has opted to lead the evaluation team and work with teachers who most needed help.² The District of Columbia Public Schools (DCPS) has taken a similar approach, appointing *master educators* with specific subject matter expertise as observers who supplement principals' observations. DCPS principals observe teachers three times each year; master educators observe them twice. These master educators follow up each observation with a half-hour feedback session in which they explain to the teacher observed why he or she received these ratings.

Redefining School Leadership

States are beginning to define the role they want principals and others to play as instructional leaders. They are building these new expectations into their evaluation systems for principals. DCPS's IMPACT system is one early example of this practice, stating performance expectations and crafting evaluation frameworks for administrators that focus on instruction. An IMPACT publication DCPS developed for assistant principals puts instruction at the center of their job. The publication tells assistant principals that their evaluation depends on their ability to "articulate a clear instructional vision," "consistently implement school-wide instructional practice," and "create opportunities

² Discussion with Katya Levitan-Reiner, May 12, 2012.

for ongoing learning and staff development.” The document asks principals to focus on human capital management, suggesting that DCPS expects them to “consistently align human resources to school needs”:

The school leader engages in effective talent management by setting high expectations, recruiting, hiring, rewarding, retaining and removing personnel, as appropriate, to ensure effective deployment of talented human capital to maximize performance and meet school goals.

The IMPACT Leadership Framework clearly spells out what high performance in this area looks like:

- The school leader carefully documents progressive discipline actions, and closely supports and tracks improvement plans for underperforming staff.
- The school leader calendars and conducts regular performance-focused conversations with all school personnel.
- The school leader manages effective, timely implementation of the IMPACT evaluation system for all school personnel.
- There is clear alignment between IMPACT scores and data (such as student achievement, facilities management and IEP timeliness), such that the performance of school personnel is accurately captured.³

This specific, detailed guidance aims to help administrators focus on instruction, including evaluation, instead of other tasks that have traditionally drawn principals away from instructional leadership.

New evaluation systems have prompted States and districts to pursue other approaches to staffing that provide administrators with time to focus on evaluation and follow-up. One promising approach is the Wallace Foundation’s school administration manager (SAM) model. It has two main components — placing in schools a school administration manager who helps shift administrative and business responsibilities away

³ “School Leader IMPACT Guidebook: The District of Columbia Public Schools Effectiveness Assessment System for School-Based Personnel,” 2011–2012, School Leaders: Assistant Principals.

from the principal and tracking the principal’s time to ensure that he or she can concentrate on teaching and learning. A study of a SAM pilot in Jefferson County, Kentucky found a doubling of student achievement gains in a one-year period. Results also suggested that principals increased the amount of time they devoted to instruction, in some cases as much as 50 percent. New York and several other States and districts have used this model to help principals assume new roles as instructional leaders.⁴

Creating Feedback Loops and Committing to Continuous Improvement

To ensure that systems are meeting the twin goals of quality and manageability, leading States and school districts through rule or policy are making commitments to continuous improvement of their evaluation systems. Virtually all of the early adopters questioned for this publication talked about the importance of this commitment. All have made regular, systematic changes to their systems in response to feedback from teachers, principals and other instructional leaders.

To this end, States and school districts can quickly implement specific feedback mechanisms, such as quarterly focus groups of teachers and observers as sources of data for continuous improvement. For this publication, DCPS and Tennessee leaders talked about the importance of regularly monitoring emails and phone calls from educators, in-person feedback, and systematic outreach to schools and LEAs. As Scott Thompson, DCPS director of teacher effectiveness strategy, underscored, leaders cannot overestimate the importance of customer service. IMPACT’s operations group includes six people whose jobs are to respond to teachers, principals, master educators and others who have questions or concerns about the system. The team maintains an email address and help line, with a 24-hour turnaround on emails.⁵ The attention the district pays to the field ensures that they are able to respond to problems with the evaluation system as they arise.

⁴ Center for the Study of Teaching and Policy at The University of Washington, Holly Holland, “Out of the Office and into the Classroom” (Seattle, Wash.: CSTP, 2008).

⁵ Discussion with Scott Thompson, May 14, 2012.

Advisory boards can play an important role in this feedback process as well, helping to vet, monitor and provide ideas about how to respond to concerns and make adjustments. In Tennessee, SEA staff members meet regularly with groups of district leaders, principals and teachers to gain ideas for improvement. Tennessee recently commissioned SCORE, a nonprofit partner, to conduct and publish an evaluation of their system — to gain outside perspective on the first year of full implementation.

In Hillsborough County, Florida, ongoing feedback has informed both the development and revision of the evaluation system to ensure it is both manageable and of high quality. Focus groups of 10 to 30 teachers convened over two years helped shape the initial design of the new system and have given way to standing teacher and principal advisory committees that have real weight in decision making. The teacher committee has had a role in major decisions, such as the choice of the Danielson framework and vetting of the value-added measures used. The committee also has weighed in on more minor but still important issues, such as the switch of the word “unsatisfactory” to “requires action” in the rating descriptors and review of new videos for accessibility. The district reinforces the standing of this committee by paying attention to their concerns and making changes to the system when warranted.

Rhode Island also has a history of using feedback to make adjustments to its system. The timeline in its Race to the Top plan and a decision to create a single statewide evaluation system designed by the State education agency meant that Rhode Island could not involve the State’s school districts in planning prior to launching the new evaluations, Rhode Island Department of Education (RIDE) officials noted.⁶ Once the new system was ready for implementation, the State listened to superintendents who urged that the launch not involve a pilot effort, but rather touch every district and every school in some way. Implementation should be gradual and more manageable, they argued. In a gradual implementation year, schools had to carry out every component of the new system, but not fully;

⁶ Discussion with Mary Ann Snider and Lisa Foehr, January 28, 2013.

evaluators conducted fewer observations and teachers developed one student learning objective, not two, for example. The approach offered an opportunity to build familiarity with specifics of the new evaluation before moving to full implementation.

Doing Thorough, Efficient Observations

In the course of implementation and especially when supported by feedback loops, States and districts can set out intentionally to streamline documents, forms and processes. Practitioners are justifiably quick to point out that complexity creates some of the most challenging aspects of new evaluation systems in terms of the time, training and skill it takes to understand and administer systems successfully.

Without strong guidance, teachers feel pressure to teach to, or check off, all the items included in lesson planning and observation frameworks every time they teach, lest an evaluator drop by unannounced and they have not demonstrated a command of several domains and multiple indicators. Principals and other observers can, in turn, spend hours preparing, completing, cataloguing and following up each time they observe a teacher. Yet, they must remember that rubrics describe excellent teaching but great teachers will not incorporate every criterion in a single lesson.

Although evaluation documents, forms and processes must be high quality and reflect a new system’s goals, some States and districts are making a concerted effort to make these elements shorter and more focused. DCPS leaders noted that they have substantially revised the original observation rubric they used in response to teacher and principal feedback, eliminating overlaps and collapsing categories that were largely redundant and moving from a rubric with 13 categories to one with nine.⁷ Denver Public Schools has responded to feedback from teachers and evaluators by reducing the number of indicators evaluators must look for and teachers must prepare for from 21 to 12. Tennessee responded to feedback from teachers and administrators by granting school districts the flexibility to reduce, for instance, the number of observations for apprentice teachers from six to three, suggesting that

⁷ Discussion with Anna Gregory and Scott Thompson, May 14, 2012.

evaluators could focus on two domains rather than one during each observation — a practice that also cuts down on the number of pre- and post-observation conferences.

Through policy and practice, States and school districts also can shorten the time necessary for formal and informal observations without sacrificing rigor. The Measures of Effective Teaching (MET) Project has provided evidence that observations need not be long. Frequent, systematic observations conducted by highly trained and certified observers yield the most effective results:

It may not be necessary for every observation to be equally long or comprehensive. Teachers who have already demonstrated basic skills could be the focus of more targeted observations aimed at higher levels of performance. In addition, systems may get a more complete picture of teacher practice if they have more frequent, shorter observations (ideally by more than one person), rather than fewer longer ones.⁸

Using Appropriate Technology

To simplify the evaluation process for teachers and leaders, leading States and districts are moving from paper and pencil systems to ones supported by technology. They are integrating what were once paper documents and forms in Web-based systems that make data collection easier and also make work flow more efficient. This has made the job of managing new evaluation processes easier for principals and teachers alike in districts such as Denver Public Schools and Green Dot Public Schools. Denver Public Schools uses a Web-based system that allows evaluators to acquire forms online, complete them there and submit them for review by the observed teacher. The teacher can then respond with his or her own comments before the principal revises and submits the final summary to the district.

⁸ Kane, T. J., and Staiger, D. O. "Gathering Feedback from Teacher Policy and Practice: Combining High-Quality Observations With Student Surveys and Achievement Gains," a MET Study policy and practice brief (Bill & Melinda Gates Foundation, 2012), 20.

Schools are adopting the use of hand-held devices for evaluation purposes as well. Some principals now use iPads to complete evaluation forms — this process allows them to enter feedback into the evaluation form on the iPads as they complete observations. Using hand-held devices saves principals a great deal of time, and it enables teachers to review feedback instantly. From the principal's perspective, this streamlined process also gives teachers time to review data from an observation before post-observation conferences.⁹

There are several vendor-created platforms that make evaluation or walk-through forms available on smartphones so that administrators can visit a classroom, perform a quick observation, record notes and scores and then upload them so they are immediately available on a Web-based platform for review by evaluators and teachers. Eliminating paperwork and processing time is a great way to reduce the burden new observation systems create.

Creating the Foundation for Success

Some States and school districts that have launched new and more rigorous evaluation systems are deploying tactics to make them more manageable but not altering fundamental assumptions or practices that are core to these new systems. Committed to continuous improvement, these States and school districts are fine-tuning their evaluation processes and protocols to ensure success.

Success, however, does not depend on these tactics alone, but also hinges on whether SEAs and LEAs lay a foundation for long term success. This foundation consists of twin commitments to the alignment of evaluation with other reforms, especially college- and career-ready standards, and educator engagement.

⁹ The Tennessee SCORE evaluation found that many principals were challenged to fulfill their traditional roles, and complete and score multiple observations. "This was highlighted as a concern across the models, with the exception of the COACH model where principals said the new evaluation system *has reduced the hours they spend on evaluations* [emphasis added], largely due to technology supports," 16.

Putting Tactics Together in Denver Public Schools

When Denver Public Schools began piloting its Leading Effective Academic Practice (LEAP) evaluation system, it had great ambitions. It introduced its “Framework for Effective Teaching,” which identified two domains (learning environment and instruction), eight expectations (classroom culture and climate, effective classroom management, standards-based goals, differentiation, high-impact instructional moves, masterful content knowledge, academic language development and 21st-century skills), and 21 indicators of performance. The LEAP evaluation system called for principals and peer observers to split equally among them four 45-60 minute formal observations, with administrators conducting additional walk-throughs and what the school district calls “partials”: 15-20 minute observations.

As the pilot year of 2011–12 wore on, the school district and the union, however, invested heavily in feedback loops. After every observation cycle, the district administered surveys. During one cycle, more than 2,000 teachers and leaders in a district serving 80,000 students completed and returned them. District leaders held focus groups, collated more than 400 emails, and met with teachers and leaders at faculty meetings to gather feedback. Teachers’ comments suggested it wasn’t fair that only the principal’s 45-60 minute observation, and not the walk-throughs and partials (many teachers had 8 or 9), contributed to their overall observation ratings. Meanwhile, the survey found that younger and probationary teachers were eager for even more feedback. It also emerged that principals — hamstrung by the requirement of devoting the same amount of time to all teachers, even the most expert instructors — could not get into the classrooms they really needed to visit.

The district and the union acted on the feedback, and the LEAP program looks very different in 2012–2013:

- The Framework for Effective Teaching is briefer. While there remain two domains, there are now only four expectations (positive classroom culture and climate, effective classroom management, masterful content delivery and high-impact instructional moves). The union and district agreed to reduce the indicators from 21 to 12.
- In compliance with State law, an administrator conducts at least one 45-60 minute observation of every teacher, and the district encourages, though will not require, the administrator to conduct at least two partials of 15-20 minutes for each teacher as well, allowing principals to exercise judgment and deploy to rooms where they are needed most.
- Peer observers conduct two 45-60 minute observations of all probationary teachers. Each year, peer observers will evaluate half the remaining teachers through two full 45-60 minute observations and observe the other half the following year. However, a principal can request that any teacher in an off year be subject to peer review.

While a little less ambitious than the original, the expectations for the revised system are still quite lofty and introduce a greater degree of human judgment about where to deploy evaluation resources.¹⁰

While a State is farther from the implementation of an evaluation system than a school district such as Denver, the State education agency (SEA) can follow a similar path, especially if it pursues a single default evaluation model. Designing and using effective feedback loops and retaining flexibility in policy can ensure that a State’s system — or local systems guided by State policies — get better over time.

¹⁰ Discussion with Jennifer Stern, July 2, 2012.

Aligning Evaluation with College- and Career-ready Standards and Instruction

Educators implementing new evaluation systems are typically implementing new standards as well. Standards for new evaluation systems measure teacher performance while college- and career-ready standards measure the performance of students. Taken as separate initiatives, implementation of both at the same time would be daunting to even the most effective teachers and principals. States are learning just now how to combine both initiatives to ensure student success and make the reforms more manageable for educators. The New York State Department of Education created www.engageny.org as the definitive resource to advance the three parts of the State's reform agenda: Common Core State Standards, data-driven instruction and school-based inquiry, and teacher and school leader effectiveness. The Website's resources detail an appropriate role for data-driven instruction and educator effectiveness in reaching college- and career-ready standards. These resources include a 27-page workbook that school districts and "network teams" — more than 800 educators statewide who are largely carrying out the State's Race to the Top plan — can use this school year to ensure the three points reinforce each other.¹¹ The workbook asks leaders to evaluate whether a district uses a common vocabulary tied to the instructional shifts required by the Common Core and interim assessments, reinforcing the central role of the student standards. It also offers "evidence guides" with tools for capturing instances of those instructional shifts in ways that relate directly to revamped teacher evaluations.¹²

In that State's largest school system, the New York City Department of Education has established "Citywide Instructional Expectations" specifically "to give New York City educators a stronger understanding of the curricular and pedagogical demands of the Common

Core."¹³ The instructional expectations describe both the Common Core standards which the city's teachers are expected to follow and the ways they are expected to implement them, including the number of Common Core-aligned literacy and math units students should experience. At the same time, New York City has been working with the unions representing its teachers and building administrators to develop a new evaluation system to comply with a 2010 State law abandoning the existing teacher evaluations that deemed teachers either satisfactory or unsatisfactory. The school district has adopted Charlotte Danielson's *Framework for Teaching* to define "the 'how' (pedagogy) that enables the 'what' of the Common Core standards."¹⁴ The district has carried out several years of pilot projects, now encompassing more than 100 schools, as a no-stakes trial of a teacher evaluation and development system based on the Danielson framework, more frequent classroom observations and the State's four-point rating scale. New York City is concentrating on only seven of the 22 competencies in the framework that it considers best aligned with the Citywide Instructional Expectations. Meanwhile, the LEA designated three competencies for special focus by all schools and included as part of the instructional expectations.

Key to the alignment of evaluation with college- and career-ready standards is a next frontier of State and district work: the alignment of observation frameworks and rubrics with these standards, including the instructional shifts they require. A future RSN publication set for release in August 2013 will address the importance of aligning expectations for students and teachers by focusing expressly on the alignment of these frameworks and rubrics with new standards. Among other important points it will make, the publication will argue that this alignment will make new evaluation systems stronger and easier to manage.

¹¹ New York State Department of Education, CCSS, APPR and DDI Workbook for Network Teams/Network Team Equivalents, October 29, 2012.

¹² New York State Department of Education, "Instructional Practice Evidence Guide for Common Core State Standards (Mathematics, Grade K-8)," November 14, 2012.

¹³ New York City Department of Education, "Implementation Guidance for the 2012-13 Citywide Instructional Expectations," June 1, 2012, 3.

¹⁴ New York City Department of Education, "Teacher Effectiveness Program: 2012-13 Handbook," (September 2012), 6.

Strengthening Evaluation by Engaging Teachers in Change

While aligning reforms is important to their success, new evaluation systems cannot succeed without the engagement of educators. Educator engagement — especially the engagement of educators at the knowledge level — is a foundation on which States and school districts will build manageable, high-quality evaluation systems.

The RSN introduced a new framework for *engaging educators* in April 2012 to help States implement Race to the Top reforms, especially new evaluation systems. The framework posits that States and school districts must engage teachers and other educators in four domains: knowledge, application, participation and leadership. It asks educators to take ownership of reforms by learning about them, applying them and then, for some, by participating in and leading the reforms.

In terms of making new evaluation systems manageable and of high quality in these early years of implementation, the application, participation and leadership domains of engagement are important but not as important as knowledge. First and foremost, teachers and principals must know what good instruction aligned to observation frameworks that are in turn aligned to new standards looks like. Secondly, they must know how the evaluation system works and how they can access support to deliver instruction aligned with instructional frameworks and rubrics.

States and districts must therefore engage *all* educators at the knowledge level.

Knowing the Basics: Good Instruction

Unless teachers and observers understand good instruction, States, districts, teachers and administrators will ultimately find new evaluation systems hard to manage and apt to fall short of success. As students need models for performance at standard, so too teachers and administrators benefit from models of effective instruction aligned with standards. They need

to see how to deliver each indicator in their framework, so that in the future they can either recognize it, deliver it or both.

Understanding this most basic need, States such as Ohio have provided video-based programs for observation rater certification to ensure that evaluators can distinguish between different levels of instructional effectiveness. Ohio has opened these training programs to teachers as well. Other States and LEAs also are finding more than initial training of observers is desirable. SCORE's evaluation in Tennessee noted, "Evaluators first complete initial training and certification, where they demonstrate the ability to rate teacher performance accurately and consistently. In each subsequent month, they participate in ongoing training and calibration to ensure strong inter-rater reliability persists over time."¹⁵

The District of Columbia Public Schools (DCPS) has amassed more than 100 video exemplars of teachers meeting the district's teaching and learning standards embedded in IMPACT. Meanwhile, this LEA is also matching the instructional shifts demanded by college- and career-ready standards with professional development offerings. To help ensure educators understand IMPACT, the district produces role-specific, easy-to-follow guidebooks to the evaluation; the teacher guidebook alone contains 74 pages of information. Denver Public Schools also anchors videos aligned with its own observation rubric on its Web-based online professional development portal.

Districts and States are learning, however, that developing high-quality videos is one thing; driving districts, teachers and schools to them is another. How States and districts encourage districts, teachers and principals to use the video banks they develop is a question that remains largely unanswered. Finding ways to utilize these videos in faculty, department, or grade-level meetings; one-on-one interactions between evaluators and teachers; or even in the

¹⁵ State Collaborative on Reforming Education (SCORE). "Supporting Effective Teaching in Tennessee: Listening and Gathering Feedback on Tennessee's Teacher Evaluation," (Nashville, Tenn.: SCORE, June 2012), 15.

private space of an individual teacher's office or home is essential to addressing the problem of how to take knowledge of what sound instructional practice looks like to scale.

Knowing the Basics: How the Evaluation System Works

Teachers and evaluators also need to know the specific details of new evaluation systems: how they will work, who the actors will be, how long it will take, what the stakes are. States and school districts might consider making clear, for example, how many observations and which type (formal and informal) evaluators will conduct, how many components of the rubric evaluators will observe when they are watching live teaching, and how the district will combine multiple observations into a single score for this component of the evaluation system. They need to know how many components and subcomponents make up the evaluation system and how district authorities will weight each of them to produce a summative rating. There should be no unanswered question and no room for rumor to take the place of fact.

Building Knowledge in the Field

Two Florida Districts in Action

Leading school districts in Race to the Top-winning States are working to build knowledge about effective instruction and evaluation systems at scale. Broward County Public Schools in Florida has worked to develop knowledge of effective teaching and its evaluation system over the last two years as it has moved to implement college- and career-ready standards, respond to the State's Race to the Top proposal and observe a Florida law changing teacher evaluation and compensation. Broward adopted the Marzano Causal Teacher Evaluation Model prior to the 2011–12 school year and provided four and a half days of professional development to its more than 1,000 administrators by November 2011, Bette Zippin, director of professional development support in the district's Office of Talent Development, recalled.¹⁶ Although the district addressed all four domains of the framework and the actual evaluation process during this extended period

of time, it spent three days focusing on one domain. Ms. Zippin reflected that the district also delivered to teachers "awareness level training" and provided them with access to an online introductory course. Teacher leaders from each school received more in-depth training to ensure that there would be a deep level of expertise at every Broward County School.

To address the instructional shifts demanded by the Common Core, Broward County created definingthecore.com, a Web portal that provides educators with access to a wide range of resources — presentations, discussion boards, lesson plans, progression guides and social media links. The district suggests that educators address particular content each month. The district attaches webinars and online self-study modules to the content. The district is also moving to create homegrown videos of teachers demonstrating the instructional shifts, moving beyond externally produced videos available thus far. "You really need to see [instruction] at various levels" of proficiency, Ms. Zippin said. "We need to see our teachers with our students."

Hillsborough County Superintendent Mary Ellen Elia has been known to say there is no such thing as too much communication. That the school district she leads has listened is evident in the wide array of communication outlets it has used in launching its new evaluation and development system and building knowledge among its corps of educators: a monthly newsletter, a quarterly magazine, a dedicated email address at which a district staff person responds to questions within 24 hours, a speakers bureau, teacher ambassadors and videos answering frequent questions or introducing topics (frequently narrated by teachers). Hillsborough is one of the largest school districts in the country. While it must rely on technology to drive communication out into the field, it still relies on face-to-face sessions when warranted. For instance, its assessment director spoke at schools across the district to explain the new calculations of value-added measures for evaluation systems. Finally, like a good teacher, Hillsborough assesses for understanding by conducting "pulse check" surveys to determine what teachers actually know after the district has communicated information.

¹⁶ Discussion with Bette Zippin, January 25, 2013.

Building knowledge of what effective practice looks like, particularly among the principals, mentors and peers who evaluate teachers, has been a focus of district professional development as well. Evaluators participate in an initial six-day training that includes video-based observation rating certification.

Applying Knowledge to Make Evaluation Systems More Manageable

To evaluators — especially principals and assistant principals — evaluation systems will seem manageable only to the degree they believe they have the knowledge to deliver on its promise, including an understanding of what effective instruction looks like. The principal also has to have the skills to apply what she knows to interactions she has with teachers. Delivering an effective conversation after an observation is one way principals apply their knowledge. In fact, this act of application is at the heart of any evaluation system, as it is in New Haven. As New Haven district administrator Katya Levitan-Reiner explained, “The design of the system is grounded in the relationship between the teacher and her manager. What’s most important is the feedback the manager provides and the relationship the manager and the teacher develop. Behavior change comes from that relationship.”¹⁷ Principals and other observers need to be able to conduct discussions that are focused on improving instructional practice and are productive for teachers. As in New Haven, Boston Public Schools is helping administrators apply what they know to after-observation discussions with teachers. In addition to giving principals a two-page reference on “look-fors” in their classroom observations, the district is beginning to train them on giving teachers professional feedback. “If we aren’t able to give meaningful feedback, we aren’t going to achieve the instructional change we want,” said Deputy Chief Academic Officer Linda Chen. “There’s an art to giving feedback so that people will respond well to it, and there’s also the content that makes the feedback specific.”¹⁸

¹⁷ Discussion with Katya Levitan-Reiner, May 12, 2012.

¹⁸ Discussion with Linda Chen, January 25, 2013.

For new evaluation systems to be manageable and successful, teachers and evaluators must know what good instruction looks like and how the system works. Evaluators must apply that knowledge when they have post-observation conferences with their teachers. More examples of engagement in the application, participation and leadership domains are available in the RSN’s publication, “Engaging Educators: A Reform Support Network Guide for States and Districts.” The importance of exemplars in delivering on the promise and sustaining new evaluation systems should not be underestimated; however, the foundation upon which States and districts build manageable and high-quality systems is a workforce in which everyone understands the reform.

States Achieving the Promise of New Evaluation Systems

As the 2012–13 school year comes to its conclusion, eight Race to the Top States and their school districts are working to achieve the promise of new evaluation systems, as they implement theirs in full. They are Arizona, Delaware, Florida, Louisiana, New York, North Carolina, Rhode Island and Tennessee. In subsequent years, more States and school districts will follow their lead. All want their new systems to result in better instructional practice and outcomes for students. Many of the early adopters mentioned in this publication have had the benefit of Race to the Top grants or, in some cases, philanthropic support. Their efforts should provide others doing the work in coming years the opportunity to adopt and adapt, rather than invent from scratch.

There is little doubt that newcomers will wrestle with what they can do to make evaluation manageable and high quality in the same way that those implementing new systems do now. States and school districts in the throes of implementing a new evaluation system naturally turn to tactics that can be executed quickly and with important effects to improve their efforts. In making educator evaluation manageable, however, the focus should remain on instruction in both the messages conveyed and in the real work being done.

In this regard, aligning reforms and engaging those working in classrooms and the principal's office is essential to success. Attention to both tactics — alignment and engagement — will ensure that more principals reach Nashville Principal Jud Haynie's conclusion: High-quality evaluation systems can in fact be manageable.

School districts will always remain closest to the actual work of evaluating teachers. Though this might be the case, States still play a vital role in ensuring that district systems are both of quality and manageable. States will remain arbiters of standards and quality and can use their policy and rule-making authority to guide school districts and promote practices that make evaluation manageable. This publication addresses a few of them: opening the door to evaluators who are not licensed administrators, and requiring an adequate but not an overwhelming number of classroom observations, for instance.

States can create feedback loops, such as those employed successfully by Tennessee, to gain greater awareness of needs and concerns in school districts. States can also support implementation to the degree resources permit by developing high-quality tools such as streamlined models of observation frameworks and rubrics aligned with college- and career-ready

standards or videos of instruction aligned with the same standards. They can provide technical assistance for single or multiple districts on many aspects of designing effective and manageable systems, including how to build knowledge in the teaching or administrative corps. The possibilities are limitless, bounded only by the creativity and ingenuity of State and district educators working hard and smart to make evaluation work for teachers so that students reach their greatest potential.

States and school districts have a big challenge ahead of them: helping teachers get better at the challenging work of teaching to new and higher expectations. The potential for fulfilling this purpose is evident in the field, as leading States and districts implement new evaluation systems that are manageable, of high quality and helpful to teachers and students in meeting these new expectations. Other States and school districts have the opportunity to build on this foundation of creativity and ingenuity, refining the practices of their predecessors and no doubt inventing new ones. Progress depends on an understanding that these new systems are not yet perfect but that with a commitment to continuous improvement and student success, new evaluation systems will achieve their potential.

References and Resources

BeAmabile, T. and S. Kramer. "The Power of Small Wins." *Harvard Business Review (HBR)*. (May 2011), retrieved from HBR Website, April 11, 2011.

Bailey, M. "New Eval System Pushes Out 34 Teachers." *New Haven Independent*, (September 13, 2011) retrieved from Website, April 18, 2012.

Boser, U. "Race To The Top: What Have We Learned from States so Far?" (Washington, D.C.: Center for American Progress, March 2012), retrieved from Center for American Progress Website, March 26, 2012.

Coggshall, A. "Toward the Effective Teaching of New College- and Career-Ready Standards: Making Professional Learning Systemic" (Washington, D.C., National Comprehensive Center for Teaching Quality, May 2012), accessed June 12, 2012.

Douglas, Emily. "The State of Teacher Evaluation," parts 1 and 2. "K-12 Talent Manager Blog," *Education Week* (Editorial Projects in Education, January, 2012), retrieved from Website April 12, 2012.

Harris Interactive. "The Metlife Survey of the American Teacher: Teachers, Parents, and the Economy." (New York, N.Y.: MetLife, March 2012), retrieved from Metlife Foundation Website April 17, 2012.

Kaine, T. and T. Staiger. (January 2012). "Gathering Feedback for Teaching: Combining High Quality Observations with Student Surveys and Achievement Gains—Policy and Practice Brief." (Measures of Effective Teaching [MET] Project, Bill & Melinda Gates Foundation, January 2012), retrieved from MET Website, May 20, 2012. MET Project Website, <http://www.metproject.org>.

National Council on Teacher Quality. "State of the States: Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies" (Washington, D.C.: National Council on Teacher Quality, October 2011). Retrieved from National Council on Teacher Quality Website, April 12, 2012.

NGA Center for Best Practices. "Preparing Principals to Evaluate Teachers" (Washington, D.C.: National Governor's Association, October 2011), retrieved from NGA Center for Best Practices Website, April 16, 2012.

"School Leader IMPACT Guidebook: The District of Columbia Public Schools Effectiveness Assessment System for School-Based Personnel: Assistant Principals," 2011-2012 (Washington, D.C.: District of Columbia Public Schools).

State Collaborative on Reforming Education (SCORE). "Supporting Effective Teaching in Tennessee: Listening and Gathering Feedback on Tennessee's Teacher Evaluation" (Nashville, Tenn.: SCORE, June 2012).

TNTP, "Teacher Evaluation 2.0." (New York, N.Y.: TNTP, (2010)), retrieved from TNTP Website, April 18, 2012.

The Wallace Foundation, "The Sam Project: Making Time for Principals to Be Instructional Leaders," (2012). See <http://www.wallacefoundation.org/Pages/SAM.aspx>.

This publication features information from public and private organizations and links to additional information created by those organizations. Inclusion of this information does not constitute an endorsement by the U.S. Department of Education of any products or services offered or views expressed, nor does the Department of Education control its accuracy, relevance, timeliness or completeness.