BUILDING TEACHER EVALUATION SYSTEMS:
Learning From Leading Efforts

MARCH 2011

INTRODUCTION

School systems across the country are working hard to fix broken teacher evaluation systems. This work offers the promise of regular, meaningful assessment of teacher performance. While this represents a significant advance, it is one part of a bigger picture: a teacher performance management system that links accountability, support, ongoing feedback, compensation, and career advancement. As school systems develop better evaluations, they will be best served if they approach this work with the other components of performance management in mind.

Moreover, systems will benefit from deliberately learning from previous attempts at innovation. This is nascent work and myriad lessons will emerge from each new initiative; the challenge is to document these lessons so school systems can determine their application in different contexts, and use them to accelerate the development and sophistication of teacher evaluation and performance management systems.

What follows is a tool for education leaders who want to learn from the experiences of the District of Columbia Public Schools[1] and the Achievement First[2] network of public charter schools. The Aspen Institute's Education & Society Program has profiled work in each of these systems. While their cultures and approaches are different, common issues emerged in both systems. By recognizing these issues and addressing them in the process of initial design and implementation, leaders in other systems can learn from and improve on prior work.

Developing and implementing a new teacher evaluation, as one component of a teacher performance management system, is deeply contextual work and no two systems will approach it the same way. The purpose of and vision for the work are informed by the system's goals, its culture and its current conditions. Sometimes the purpose is to dramatically alter the current culture and conditions; other times the purpose is to strengthen budding or well-established practices. Whatever the case, the challenge school systems face is to simultaneously focus on the details of the evaluation system they are trying to build while keeping an eye on the bigger picture of performance management and how it will support district-wide improvement in instruction and student achievement.

Better systems – and better results – will be realized if these questions are considered and addressed prior to design and implementation of new teacher evaluation systems.

[1] Rachel E. Curtis, "District of Columbia Public Schools: Defining Instructional Expectations and Aligning Accountability and Support," March 2011. Available at http://www.aspeninstitute.org/education.

[2] Rachel E. Curtis, "Achievement First: Developing a Teacher Performance Management System That Recognizes Excellence," March 2011. Available at http://www.aspensinstitute.org/education.

AS SCHOOL SYSTEMS ENGAGE IN THIS WORK, HERE ARE FIVE QUESTIONS THAT ARE WORTH ASKING TO GUIDE THEIR EFFORTS:

1. What level of specificity and conformity are we trying to create through our teaching standards?

2. How do we balance high-stakes accountability with ongoing support and feedback that are essential for improvement?

3. How will teachers be engaged to build ownership of teacher evaluation and performance management?

4. How will the system's teaching and learning infrastructure need to evolve to support implementation of the teaching standards?

5. How will the system need to function differently to implement a robust performance management system and how will the district build this capacity?

WHAT LEVEL OF SPECIFICITY AND CONFORMITY ARE WE TRYING TO CREATE THROUGH OUR TEACHING STANDARDS?

1

How school systems define teaching standards falls along a continuum. Achievement First's (AF) approach is at one end of the continuum, explicitly articulating how teachers will do specific things (e.g. use exit ticket to check understanding and assess learning at the end of each lesson; structure each lesson with a mini-lesson, guided practice, and independent practice).

The District of Columbia Public Schools' (DCPS) standards sit at the middle of the continuum articulating general expectations about what teachers will do (e.g. engage students; check for understanding) without mandating specific instructional moves. The Danielson Framework sits towards the opposite end of the continuum from AF, talking in holistic ways about instructional practice (e.g. teacher demonstrates moderate flexibility and responsiveness to student questions, needs and interests during a lesson and seeks to ensure the success of all students). This encourages teachers and the evaluators who use the standards to construct their own meaning of them. Any of these approaches can work, but it is important to understand the different implications of each.

The more specific and discrete the standards are, the easier it is to train teachers on them and measure their performance against them. This approach may make it easier and faster to realize greater consistency in instructional practice. But it runs the risk of reducing teaching – a complex task – into a series of moves which has the potential to limit teachers' growth and development. This approach can also limit teacher engagement and ownership of the standards because there is little room for teachers to make meaning of them. DCPS, whose standards sit in the middle of the discrete to holistic continuum, quickly realized that when its standards rubric was too specific about what should be happening in the classroom (e.g. four out of four students surveyed can articulate the lesson objective; 2/3 of students engaged and on task), evaluators became too focused on counting hands in the air and trying to discern what was an appropriate articulation of the lesson objective from a third grader. This distracted evaluators from making holistic observations about the quality of instruction.

In contrast, more general statements about instruction require teachers and their supervisors to interpret the standards. This requires significant professional development to build understanding of the standards and the instructional practices that are aligned to them as well as conversations between teachers and their supervisors as they talk through observations and make sense of them in the context of the standards. Both of these things take significant time which is generally in short supply in schools. Yet it is that time spent making meaning of them that can lead to a greater sense of teacher and evaluator ownership of the standards.

Opting for broad standards requires figuring out how to make them concrete and specific, while adopting very discrete standards requires a strategy to generalize from them to a theory of effective instruction.

2

HOW DO WE BALANCE HIGH-STAKES ACCOUNTABILITY WITH ONGOING SUPPORT AND FEEDBACK THAT ARE ESSENTIAL FOR IMPROVEMENT?

Given the long history of weak teacher evaluations, developing and implementing a high-quality evaluation system is a tremendous step forward. Creating such a system is a complex act that requires an infrastructure of standards, professional development, trained evaluators, information management systems, and quality control processes that most school systems must build. When all of this is in place the system will tell teachers how they are performing and give the school district important information about teacher performance that can then drive training and development, rewards and consequences.

For most school systems, getting such a system up and running is initially all consuming. As demanding as it is, the effort will be both necessary and insufficient. Evaluation ratings don't provide guidance to teachers on how to improve their practice. Yet research tells us that teachers improve when they receive regular, specific feedback on their practice; the opportunity to reflect on it; and support to improve areas of weakness.[3]

Strong performance management systems integrate evaluation with supervision that supports growth and development and are akin to student assessment systems that provide formative, interim and summative data. Achievement First's performance management system uses a mix of formative, interim and summative assessments. Teachers are observed weekly or bi-weekly and given timely, specific feedback on their practice by a coach whose job it is to support their growth and development. Teachers debrief the observations with their coach and this team defines goals for the teacher's growth that they then work on together. This formative cycle of observations, feedback, goal setting and coaching continues throughout the year punctuated by a mid-year interim assessment and annual summative evaluation of the teacher's practice.

An evaluation cycle that includes multiple observations over the course of the school year each of which is scored and then factored into a cumulative assessment (DC's system is an example of this) shares similarities with interim assessments students take throughout the year that indicate where they are in their learning relative to the standards. As with interim assessments of students, the value isn't just in collecting the information, but in tapping its diagnostic potential to guide reflection and adjustments to practice.

In contrast, a once-a-year observation that culminates in a written evaluation is a bit like the annual state assessment; it tells us something important and limited about a teacher's performance at a moment in time – a moment in time that has already passed. The lessons learned about the limitations of summative student assessments – particularly with regard to informing practice – hold true for a summative approach to teacher evaluation.

Figuring out when and how often to collect teacher performance data and what to do with it is a critical step in developing a teacher evaluation system. Using the data to hold teachers accountable supports a strong evaluation system. Being equally committed to using the data to support teachers to improve creates the conditions for a performance management system. The design of evaluation systems should be carefully considered to ensure they reflect system leaders' beliefs about their responsibility relative to holding teachers accountable for performance and supporting them to improve.

[3]Eleanor Drago-Severson, *Helping Teachers Learn,* Corwin Press, 2004.

HOW WILL TEACHERS BE ENGAGED TO BUILD OWNERSHIP OF TEACHER EVALUATION AND PERFORMANCE MANAGEMENT?

3

Building teacher ownership of standards and everything that is aligned to them is essential to the success of a new evaluation system. Ownership impacts the credibility of the effort in the hearts and minds of teachers, and affects whether the standards are seen as unreasonable expectations foisted upon them by people far away from the classroom or a powerful vision of teaching excellence.

There are many ways to engage teachers in this work. At the standards development stage, involving teachers in writing the standards and sharing the standards at different stages of development for feedback are powerful ways to both engage teachers and strengthen the standards. Credible teacher engagement efforts are substantive and meaningful rather than superficial and perfunctory, and teachers' voices are given as much credence as those of administrators and outside consultants.

Once the standards are defined, piloting them with a group of teachers who provide feedback and serve as critical friends to the system allows for a sheltered implementation of the standards with a built-in process for learning and improvement. The roles teachers play in supporting standards implementation can further develop ownership. Engaging teachers whose work embodies the standards to support their colleagues in implementation broadens ownership. Celebrating teachers whose practice reflects the highest possible performance relative to the standards is another way to broaden ownership.

The decision to have teachers step out of the classroom and assume the role of evaluating other teachers (in the DCPS case they did this from an administrative position) fundamentally changes the evaluation dynamic. Something that was historically quietly shared between teachers and their principal (and perhaps another school administrator) is shared more broadly and more publically. Also, because teachers who take on this evaluative role are chosen, in part, for their instructional expertise, the opportunity for teachers observed to get specific feedback on their practice is increased.

Communication is as important as engagement in building ownership. Trust is essential to implementing anything new effectively and to building a culture of shared ownership. Building trust depends on clear, consistent communication.[4] In systems with weak cultures, an information vacuum is often filled by people's worst fears. Given the high-stakes nature of new teaching standards and evaluation, this is especially likely without open lines of communication. Frequent, clear communication provided at every level of the organization is critical – in-person, in writing and online. When teachers, principals, and central office staff repeatedly hear the same clear message from the senior leaders of the system, they begin to give more credence to the initiative. Having a high-touch approach with broad outreach is helpful as messages inevitably get distorted as they work their way through the organization.

Encouraging two-way communication, where feedback from teachers and principals in the throes of implementation comes back to the system and informs refinements, both builds ownership and strengthens the evaluation system. This can be done through surveys, focus groups and ongoing work groups. The level of trust and collaboration should inform how information is collected. In some systems an anonymous survey is required to get meaningful feedback, while teachers in other systems are very comfortable speaking about their concerns in a public forum.

4

HOW WILL THE SYSTEM'S TEACHING AND LEARNING INFRASTRUCTURE NEED TO EVOLVE TO SUPPORT IMPLEMENTATION OF THE TEACHING STANDARDS?

To maximize effect, new teacher evaluation needs to be inextricably linked to the system's teaching and learning work. This partnership is critical to ensure teachers have access to professional development and the curricular, instructional and assessment resources they need to help them meet the teaching standards.

Professional development needs to be provided at three levels to support full implementation: 1) awareness; 2) deepening pedagogical expertise; and 3) applying the standards in content-area teaching. These three levels tend to be addressed sequentially. At first, professional development focuses on making teachers and their evaluators aware of the standards and what they look like in action. Once an awareness level has been developed, teachers need opportunities to deepen their understanding and develop strategies to adjust their instruction to reflect the standards. Finally, once teachers understand the standards and have begun to use them to guide their practice, the next phase of the work can focus on how the standards are reflected in content-area teaching.

Standards tend to be content-neutral, focused on pedagogy generally so that they are relevant to all teachers, but this only gets us so far. To get the greatest improvements in teaching practices, standards need to be defined in the context of teaching specific content – i.e., what does checking for understanding and clearing up student confusion look like in a second-grade guided reading lesson as compared to a high-school biology class that includes labs.

To do all of this, school system's professional development delivery systems – new teacher induction, common planning time, instructional coaching, professional learning communities, workshops and school-based professional development – all need to be aligned with helping teachers meet the standards. The teachings standards must become the spine running through all teacher development activities. If every office and initiative uses its own theory about good instruction and its own vocabulary, teachers get inconsistent and incoherent messages, and implementation of the standards is diluted. For example, DCPS realized early in its implementation that school-based, instructional coaches were not integrated deeply enough into implementation of the standards. As a result, coaching wasn't reinforcing the standards and was sometimes providing mixed messages to teachers about the district's expectations.

It is almost inevitable that the implementation of teaching standards will surface points of misalignment of the teaching and learning infrastructure. Systems need to examine patterns of weak teacher performance to assess when these are the result of the system not providing critical support. Widespread weaknesses in teacher planning and lessons that lack rigor often indicate a system's lack of clear curriculum on which teachers can organize their planning and/or a common approach to lesson planning. Weak use of student assessment data to measure learning and inform instruction is often the result of a system not having assessments available that give teachers real-time information they can use to guide their instruction. In other systems, the assessments are in place, but there is not a focus on teaching teachers how to analyze the data and use it to inform their instruction.

Anticipating these challenges, developing a strategy to address them, and discussing what it is fair to hold teachers accountable for and what school systems must be held accountable for providing to teachers are critical to building a system with integrity.

HOW WILL THE SYSTEM NEED TO FUNCTION DIFFERENTLY TO IMPLEMENT A ROBUST PERFORMANCE MANAGEMENT SYSTEM AND HOW WILL IT BUILD THIS CAPACITY?

5

As important as it is to have a vision for the performance management work, it is equally important to think developmentally about how to build the capacity to realize that vision. A vision of teaching standards guiding teacher hiring, induction, tenure, support, evaluation, compensation and career advancement requires a retooling of all these functions and implicates multiple central office departments (e.g., human resources, information management, teaching and learning). It is unlikely that the system has the capacity to undertake all this work at once.

Most school systems are choosing teacher evaluation as the entry point into performance management. This raises an important question: How will the foundation created by stronger evaluations support the broader vision of a performance management system?

Each school system's particular context will shape its priorities. A system that is hiring a significant number of new teachers may decide to align its induction support and tenure review systems to the standards early on. For a system focused on building teacher leadership, identifying opportunities and incentives for teachers whose performance relative to the standards is exemplary may be a top priority. A system that has a strong culture of teacher development and support or wants to develop one may focus all year-one implementation on support, waiting until the second year to introduce the accountability dimension. Conversely, a system suffering from weak practice and a low level of accountability may decide that the standards and an evaluation that measures performance against them must be introduced simultaneously.

Regardless of where a school system chooses to start implementation of the standards as part of a performance management system, collaboration across departments will be required. Structures and systems can be put in place that can either support or impede collaboration. One pivotal decision is where the work of performance management sits within the system's organizational structure. Does it sit in human resources, teaching and learning or a new division of human capital that reports directly to the superintendent?

If the support and accountability functions are separated organizationally, the challenge is how to incentivize and build expectation for a high level of collaboration across the departments to ensure tight alignment of efforts and a coherent experience for teachers. This raises questions about cross-functional work such as: Is the evaluation division responsible for measuring every teacher's use of student assessments to inform instruction only for evaluative purposes, or are they also responsible for analyzing those data with the teaching and learning division to identify implications for professional development? Anticipating the organizational demands of this work and structuring the work to maximize success and force new ways of working provides tremendous opportunities.

It is very clear from AF's and DCPS's experiences that successful implementation of a new teacher evaluation system and using it to anchor a performance management system requires continuous learning and refinement. The clearer systems can be about what they need to learn and how they will learn it, the better positioned they are to ensure the quality of their system. Some learning priorities are easy to identify from the outset (e.g. teacher attitudes and experiences, calibration of evaluators' scores, efficacy of professional develop-

ment) so the system can plan to collect and analyze data and use it to inform improvement. Other issues may surface during implementation and systems need to have a way to identify and respond to them.

From the beginning, DCPS knew that the integrity of its evaluation system hinged on calibrating evaluators' rating to ensure that teachers would be scored similarly regardless of who observed them. At the end of each observation cycle throughout the year, DCPS analyzed the data of all the evaluators to discern where differences existed, to develop hypotheses to explain the difference, and then strategies to eliminate them. DCPS also anticipated that the standards and rubric would need to be refined at the end of year one because the system would have learned so much that could be used to strengthen them. The system communicated this early on in the first year of implementation and repeatedly throughout the year. Given this commitment, the question the system had to answer was: how will it collect and analyze the information needed to inform refinements and who will be involved in this process?

Being clear about the system's intention to learn from its work and improve sends several powerful messages that support both the success of the evaluation effort and a strong organizational culture. It signals to teachers and evaluators that their perspective is valued, which helps build ownership. It clearly communicates an understanding that complex work cannot be done perfectly from the start so it must be undertaken with a commitment to learning and improving. This message has the potential to encourage individual departments and individual staff members to take a similar learning stance in their work.

CONCLUSION

Putting a new teacher evaluation system in place is complex work with a long reach. To ensure a new evaluation system supports the larger goal of improving teaching and learning, school systems need to simultaneously delve deeply into the specifics and step back far enough to see the long view. It is true that establishing a meaningful teacher evaluation system represents significant progress for many school systems and the sector as a whole. Yet, if this work is approached systemically and strategically, it has the potential to dramatically change how teachers think about their practice and their role in the system, how school systems function, and what they do to support instructional excellence. Getting this right is essential to driving improvement in instruction and student learning.

THE ASPEN INSTITUTE

THE MISSION OF THE ASPEN INSTITUTE IS TO FOSTER ENLIGHTENED LEADERSHIP AND OPEN-MINDED DIALOGUE. THROUGH SEMINARS, POLICY PROGRAMS, CONFERENCES, AND LEADERSHIP DEVELOPMENT INITIATIVES, THE INSTITUTE AND ITS INTERNATIONAL PARTNERS SEEK TO PROMOTE NONPARTISAN INQUIRY AND AN APPRECIATION FOR TIME-LESS VALUES. THE INSTITUTE IS HEADQUARTERED IN WASHINGTON, D.C., AND HAS CAMPUSES IN ASPEN, COLORADO, AND ON THE WYE RIVER ON MARYLAND'S EASTERN SHORE. ITS INTERNATIONAL NETWORK INCLUDES PARTNER ASPEN INSTITUTES IN BERLIN, ROME, LYON, TOKYO, AND NEW DELHI, AND LEADERSHIP PROGRAMS IN AFRICA AND CENTRAL AMERICA.

THE ASPEN INSTITUTE
EDUCATION & SOCIETY PROGRAM

THE EDUCATION AND SOCIETY PROGRAM PROVIDES AN INFORMED AND NEUTRAL FORUM FOR EDUCATION PRACTITIONERS, RESEARCHERS, AND POLICY LEADERS TO ENGAGE IN FOCUSED DIALOGUE REGARDING THEIR EFFORTS TO IMPROVE STUDENT ACHIEVEMENT, AND TO CONSIDER HOW PUBLIC POLICY CHANGES CAN AFFECT PROGRESS. THROUGH OUR MEETINGS, ANALYSIS, COMMISSIONED WORK, AND STRUCTURED NETWORKS OF POLICY-MAKERS AND PRACTITIONERS, THE PROGRAM FOR NEARLY 30 YEARS HAS DEVELOPED INTEL-LECTUAL FRAME-WORKS ON CRITICAL EDUCATION ISSUES THAT ASSIST FEDERAL, STATE, AND LOCAL POLICYMAKERS WORKING TO IMPROVE AMERICAN EDUCATION.