



MARCH 2016

Teacher Evaluation *and* Support Systems

A Roadmap for
Improvement





THE ASPEN INSTITUTE

EDUCATION & SOCIETY PROGRAM

The **Aspen Education & Society Program** improves public education by inspiring, informing, and influencing education leaders across policy and practice, with an emphasis on achieving equity for traditionally underserved students. For more information, visit www.aspeninstitute.org/education and www.aspendrl.org.



THE ASPEN INSTITUTE

The **Aspen Institute** is an educational and policy studies organization based in Washington, DC. Its mission is to foster leadership based on enduring values and to provide a nonpartisan venue for dealing with critical issues. The Institute is based in Washington, DC; Aspen, Colorado; and on the Wye River on Maryland's Eastern Shore. It also has offices in New York City and an international network of partners.

Copyright © 2016 by The Aspen Institute

The Aspen Institute • One Dupont Circle, N.W., Suite 700 • Washington, DC 20036

Published in the United States of America in 2016 by The Aspen Institute

All rights reserved

Preface

Decades of experience and research tell us that teachers matter more to student achievement than any other aspect of schooling. States have an important role to play in making sure teachers have the support they need to excel in their increasingly complex roles. A strong teacher support and evaluation system can help states in building, developing, and retaining an educator workforce that can realize the promise of public education for all students.

As the executive director of the Council of Chief State School Officers (CCSSO), I applaud The Aspen Institute for this comprehensive set of recommendations on how states can refocus and refine their teacher evaluation systems to support teachers' professional growth and development. *Teacher Evaluation and Support Systems: A Roadmap for Improvement* provides clear recommendations on how states can continue to improve the support and evaluation systems they have in place.

Our goal at CCSSO is to make sure every public school student — regardless of background — graduates prepared for college, careers, and life. This work begins with a highly effective teacher in every classroom, and teacher support and evaluation systems are critical in achieving this goal. While many states have designed and implemented new teacher evaluation systems, the passage of the Every Student Succeeds Act gives each state a valuable opportunity to reexamine these systems and policies to determine what they will keep and what they will change to continuously improve those systems as the needs of practitioners and students change. To support states in that work, CCSSO has created *Principles for Teacher Support and Evaluation Systems*.

The recommendations in *Teacher Evaluation and Support Systems: A Roadmap for Improvement* complement CCSSO's *Principles* and offer concrete recommendations and examples for states to consider as they take a new look at their evaluation systems. We believe every state should review these recommendations and consider them, along with CCSSO's *Principles*, as they continue their work to improve evaluation and support systems for all educators and to ensure every child has access to a highly effective teacher.



Chris Minnich
Executive Director
Council of Chief State School Officers

Recommendations to Improve Teacher Evaluation and Support Systems

- 1 Prioritize principal and evaluator training and certification with a focus on professional growth
- 2 Differentiate evaluation and support based on teachers' experience and past performance
- 3 Allow teachers and observers to collaborate on areas of focus
- 4 Allow for local discretion in accounting for student learning
- 5 Respect the limitations of value-added data
- 6 Support locally developed measures while pursuing improvements in their creation and use
- 7 Make sure all important aspects of teaching performance are valued in evaluations
- 8 Engage teachers in improving teacher evaluation systems
- 9 Develop measures for testing the integrity of evaluation system design and implementation
- 10 Tell stories that go beyond performance ratings

Introduction

Policymakers and educators have invested significant time, effort, and political capital in developing policies and processes for evaluating teacher performance.

These activities were driven by a recognition of the powerful impact teachers have on learning and an understanding that traditional evaluation systems were not helping to improve teachers' practice. As with any undertaking of this importance and complexity, it is essential to revisit assumptions, learn what is working and what is not going well, and continuously improve the system. States, districts, evaluators, and teachers can take their significant implementation experience and move

beyond a focus on compliance and accountability functions toward the use of evaluations as a resource for teacher growth and development.

The passage of the federal Every Student Succeeds Act (ESSA) in December 2015 and the accompanying end of federal policies and requirements regarding teacher evaluation provide a natural inflection point for states to review existing systems and recommit to the primary purpose of evaluation: to support teacher growth and development and to act as but one component within robust systems of talent management and instructional improvement. States and districts are at different stages of implementing their teacher evaluation systems, and some have already modified their systems or delayed some consequences based on practical or political considerations.¹ Although some states are likely to keep their teacher evaluation systems unchanged, others are reconsidering teacher evaluation policies in light of ESSA.² Whichever path leaders choose to follow, states and districts should focus on teacher growth and continuously improve the system. While much of the debate around evaluation has been polarizing, the recommendations in this report offer a middle path, encouraging state leaders to revisit assumptions, examine what's working (and for whom), and acknowledge some missteps while still making good on the promise of evaluation and support systems overall.

The recent experiences of states and districts suggest that key issues in the implementation of teacher evaluation systems include:

- The proper division of responsibility between State Education Agencies (SEAs) and Local Education Agencies (LEAs) for designing evaluations.³
- The training and capacity of principals and other classroom observers and evaluators.
- The use of meaningful measures of student growth.
- Personalization of the evaluation and support process.
- The tension between state formulas and professional judgment in assessing a broad body of evidence related to teacher performance and practice.
- The degree of teachers' involvement in decision-making about which aspects of their practice should be the focus of evaluations and in providing feedback on whether evaluations are helping improve practice.

This report examines these issues, highlights states and districts that are leading the way, and outlines 10 recommendations to improve teacher evaluation systems. Some of the proposed strategies may be easy changes for states and districts to make, while others will probably be more challenging. No state is likely to be interested in adopting all of these recommendations, but state leaders can use this list to provoke dialogue about the purposes of the evaluation process and opportunities to achieve the overall promise of educator effectiveness systems.

Good to Great and Great to Excellent

As states and districts revisit their evaluation policies, the process will necessarily be informed by how well they implemented their evaluation systems and by the political and historical context, but all will find useful support in the framework developed by McKinsey & Company based on an examination of how the world's best school systems continuously improved.⁴

require more focus on fidelity to models, fewer options for local discretion, and tighter oversight by the state or district. As reforms progress and systems advance from good to great and from great to exemplary, more discretion is afforded for professional judgment and local decision-making.

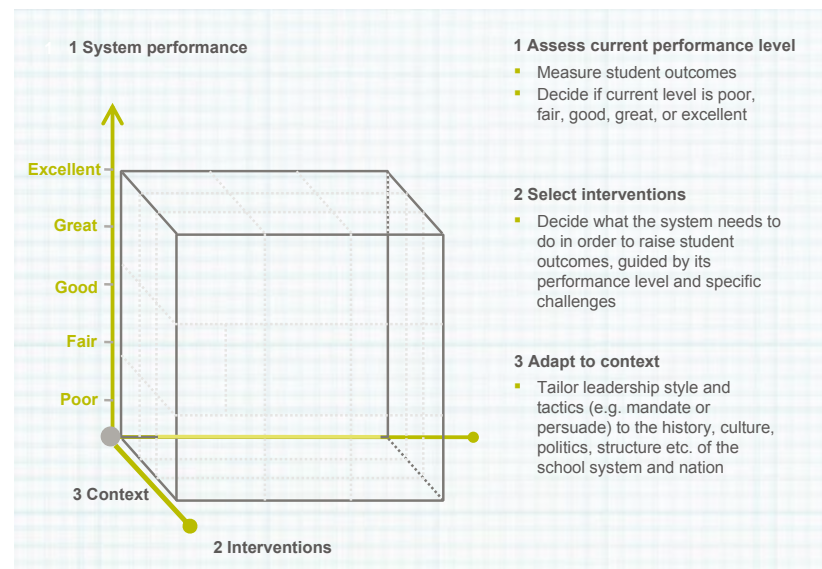
The analogy to the improvement of evaluation processes is clear. When education leaders and policymakers are moving their teacher evaluation systems from fair to good, they need to require adherence to a well-defined systemwide model. Schools and districts should be given fewer options and implementation should be closely monitored. However, as educators and schools master the evaluation process and the system strengthens its overall infrastructure, they should be given greater discretion in exercising their professional judgment.

Many states and districts have already made notable progress in establishing more rigorous evaluation systems and, as a result, are positioned to support even more significant gains in teaching and learning. For example, the **District of Columbia Public Schools'** (DCPS) IMPACT teacher evaluation system was initially put in place when instructional quality was a major problem and district leaders needed to

The McKinsey framework authors found that all of the school systems they studied focused on the same six strategies, regardless of how far along they were on their path toward excellence.

However, the manner in which school systems pursue those strategies changes as their efforts and procedures become more fully developed. When leaders are moving a system from poor to fair and from fair to good (as has arguably been the case for most teacher evaluation systems in the United States), policies often

McKinsey & Company framework: Three dimensions a system leader must integrate when crafting and implementing an improvement journey



Source: Mona Mourshed, Chinezi Chijoke, and Michael Barber, "How the World's Most Improved School Systems Keep Getting Better," McKinsey & Company, November 2010, available from http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/social%20sector/pdfs/how-the-worlds-most-improved-school-systems-keep-getting-better_download-version_final.ashx.

impose clarity and consistency. Dozens of master educators were trained and supervised centrally, and the district saw some notable success in improving instruction and student achievement.⁵ In the 2016-17 school year, DCPS principals will reclaim more responsibility for teacher evaluation, and resources will be reallocated away from master educators and toward teacher leaders, team collaboration, and expert coaching that is not connected solely to evaluation.⁶ Once the fundamentals of an evaluation system are in place, going from good to great may require more autonomy for school leadership teams to meet the specific needs of their students and teachers.

This raises important questions for state leaders and policymakers, including “Where is my system on the continuum from poor to exemplary, and how can policies be calibrated correctly?” and “How can states tightly manage systems going from poor to fair to good, while allowing greater flexibility for higher-functioning systems to move from good to great to exemplary?”

THE ROLE OF STATES

States have a great deal of influence over the preparation, development, and retention of an effective educator workforce. With this in mind, the Council of Chief State School Officers (CCSSO) in March 2016 issued a set of guiding principles for the design of teacher support and evaluation processes, which can help states keep the goals of evaluation front and center. The 10 recommendations in this paper are intended to be a complement to the CCSSO principles and provide guidance on concrete next steps for states.

CCSSO'S PRINCIPLES FOR TEACHER SUPPORT & EVALUATION SYSTEMS⁷

1. Integrate teacher support and evaluation into broader efforts to develop teaching practice and improve student learning.
2. Drive continuous improvement of teaching.
3. Ensure the system is fair, credible, and transparent.

STATE AND LOCAL EDUCATION AGENCIES

States and districts each have important roles in the design and implementation of teacher evaluation systems. Although the division of labor will vary across states, all state education agency (SEA) leaders need to collaborate with their counterparts at the local level and other stakeholders to explicitly allocate responsibility. No single approach to defining state and local roles will work for every context, but however officials choose to go about it, their decisions should be deliberate, strategic, and clearly communicated to all system leaders.

Some roles clearly should be shouldered at the state level (e.g., calculating value-added or growth data in tested grades and subjects, so the reference group is statewide rather than district-limited), other decisions are better managed locally, and some issues demand shared responsibility. For example, some states might establish certain requirements statewide and identify other elements that must be in place but invite local decisions on the details, while other aspects might be entirely at the discretion of local education agencies (LEAs).

Similarly, while not displacing local authority, states might set quality assurance standards, highlight exemplars, and create opportunities for networks to accelerate learning.

Among the recommendations in this paper, some might be more appropriate for LEA consideration based on the division of labor within the state. Further, the division of labor that works with one LEA may not be right for all those in the state. States might assess the appropriate roles for their LEAs in several ways, including:

- **District capacity:** Some LEAs can bring as much or more personnel and expertise as the SEA to critical design decisions, while other, often smaller districts are more reliant on state capacity.
- **Quality of prior implementation:** SEAs can start with tight rules but allow greater flexibility to systems that demonstrate quality implementation. This approach also might represent a strategy for fostering innovation at the local level.

Recommendations

1 Prioritize principal and evaluator training and certification with a focus on professional growth

In 2015, 17 states required principals and other educators to pass a test to become certified evaluators of teachers' performance.⁸ However, most of the certification tests only asked principals and other evaluators to accurately score a videotaped lesson. This approach to certification sends the message that the main purpose of evaluation is to rate teachers, rather than to support their professional growth. Moreover, this practice focuses narrowly on the act of teaching rather than considering the full range of activities that are part of a teacher's professional

responsibilities, such as lesson planning, analyzing student data, building school culture, and collaborating with colleagues. To become certified evaluators, leaders should be required to demonstrate that they can:

- Make concrete, actionable recommendations to help teachers improve their practice.
- Provide constructive feedback to teachers based on observations, student work and assessment data, survey responses, and other elements of the evaluation process.
- Use the results of both individual and schoolwide evaluations to design professional learning activities.

Research from the Consortium on Chicago School Research shows that the adaptive aspects of teacher evaluation—such as having the skill and will to provide meaningful, constructive feedback to a colleague—are more challenging for principals than the technical aspects of rater accuracy and reliability.⁹ The certification of evaluators should reinforce the idea that improving instruction is an essential, primary purpose of the evaluation system, as opposed to merely rating accurately against an observation rubric.

For example, the statewide **Tennessee** Educator Acceleration Model (TEAM) uses the TAP (The System for Teacher and Student Advancement) Teaching Standards—which are provided to districts by the National Institute for Excellence in Teaching—as the basis for classroom observations.¹⁰ To be certified as a TEAM evaluator, principals and other educators must show they can score a video lesson accurately, develop a detailed improvement plan, and facilitate an effective post-evaluation conference. The improvement plan is judged by an external expert based on a rubric that considers the evaluator's ability to provide feedback and support, focusing on specific aspects of instructional and professional growth that were identified in the observation.

States and districts have struggled to equip principals to evaluate and develop teachers, and that has exposed gaps in policies and practices for training, licensing, supervising, supporting, and evaluating principals as well. Many states already have administrator evaluation systems in place, and some, like **Massachusetts**, are incorporating the skills needed for effective teacher evaluation into principal certification programs and requiring performance assessments that include delivering feedback as part of the requirements for initial licensure.¹¹ School systems need to provide ongoing training and support to principals and other leaders to meet the evolving demands of their roles as instructional leaders.

2 Differentiate evaluation and support based on teachers' experience and past performance

The first few years of teaching demand intensive growth as new teachers learn to focus on classroom and behavior management and routines, as well as pedagogical approaches, lesson planning, grading, building relationships with peers, and time management—often with little support or time for reflection.

New teachers can benefit from a focused evaluation and support process with evaluators who concentrate on the skills and behaviors most associated with early success, such as planning, classroom environment, and instructional techniques. Evaluators should emphasize “developing” or “satisfactory” performance levels rather than the indicators associated with “exemplary” performance and provide novices with the clear feedback and explicit guidance they need to become more effective. How inexperienced teachers

respond to such direction will give evaluators a sense of how deeply committed the teachers are to improving their practice, as well as whether they are making adequate progress.

In **Massachusetts**, for example, officials assume that few new educators will demonstrate proficient practice on all standards in their first years on the job, so new teachers are automatically placed on a Developing Educator Plan, regardless of their performance or their students' test scores.¹² As part of the process, novice teachers are observed more frequently and engage in focused goal-setting on a truncated timeline compared to their more experienced peers. According to Heather Peske, Associate Commissioner for Educator Effectiveness, this policy shelters novice teachers from high-stakes, negative consequences while they are learning the job and helps the state and districts transition new talent into the system.¹³ The Developing Educator Plan can also be used with experienced, effective teachers who take on new and more challenging assignments, insulating them from high stakes and encouraging teachers to step up to new challenges. Teachers with a Developing Educator Plan are expected to achieve proficiency on all evaluation components within a designated time frame and to use the plans they co-develop with their evaluators to determine how they will meet those expectations. The intent is not to send the message that developing teachers can't or won't excel, but to acknowledge the complexity and challenges of the first few years of teaching and of tackling new assignments.

Evaluations of veteran teachers who have previously been rated below effective should be even more prescriptive and include more frequent observations and shorter-term development goals. This level of supervision, however, would require principals to spend a great deal of time on a small number of teachers, taking them away from other duties. To shift some of the burden from principals, districts may want to assign responsibility for working with struggling performers to specially trained personnel who can help teachers improve quickly or determine that they should be removed from the classroom.

Some systems, such as **Cincinnati Public Schools** and **Montgomery County, MD**, use a Peer Assistance and Review (PAR) process in which teacher leaders, also called consulting teachers, coach struggling educators and recommend dismissal when appropriate.¹⁴ Consulting teachers are selected via a rigorous screening process that considers their classroom practice, and they are trained to provide support to struggling or novice teachers. A similar program in **Tennessee** that uses evaluation data to pair poorly rated teachers with peers who have demonstrated strong performance in the same area has proven to be an effective strategy for improving instruction.¹⁵

Effective teachers, on the other hand, should be allowed to customize their evaluations in partnership with their principals or with teams of teachers in their schools or districts. For instance, they might be subject to fewer formal observations than teachers who are still mastering their craft, but they could still benefit from informal observations that help further hone their practice and target specific areas for growth. It is counterproductive, however, to make effective teachers feel that they are being neglected by giving them less attention or support. These are teachers who are making significant contributions, and schools systems need for them to feel valued and supported.¹⁶ To make evaluation processes more manageable, some systems are simply cutting the number of required observations for effective teachers, reinforcing the notion that evaluation is primarily for identifying and removing low performers. Differentiation should not be a matter of cost savings or efficiency, but rather a recognition of achievement and an investment in advancing performance to the next level.

Just as evaluation ratings can be useful in differentiating subsequent evaluation processes, they should also be used to differentiate support; highly effective teachers should be given more flexibility to identify their own professional growth plans based on the needs of their students, as well as objectives decided upon by their grade level or subject matter teams.

In **Massachusetts**, teachers rated Proficient or Exemplary during their previous evaluation cycle are placed on self-directed growth plans, which give them opportunities to work toward longer-term goals and ensure that their professional growth needs are met. Because these educators develop their own plans, evaluators are free to focus more of their time on new and struggling teachers.¹⁷ Similarly, the **Achievement First** Charter Management

Differentiation should be a recognition of achievement and an investment in advancing performance to the next level.

Organization (CMO) offers funding and supplemental opportunities for self-directed professional development to teachers who attain advanced stages (senior teacher and master teacher) on the CMO's career pathway.¹⁸

Highly effective teachers should also have opportunities to continue their own growth by developing other teachers, both within and beyond the evaluation process. In **Shanghai, China**, for example, as teachers ascend the career ladder, their evaluations hinge on their ability to lead learning among their colleagues.¹⁹ These teachers can assume the roles of adjunct faculty to train new teachers or serve as evaluators and mentors to less-experienced colleagues.

Each of these roles extends the reach and impact of the best teachers. To maximize effectiveness of teacher leadership programs, however, policies and structures must be formalized (i.e., with clear pathways to these roles; transparent selection criteria; explicit links to teachers' evaluation, growth, and development; and consistency across peers, schools, and districts) instead of the ad hoc leadership roles that are more typical today.²⁰

3 Allow teachers and observers to collaborate on areas of focus

A comprehensive teaching framework is essential for defining high-quality teaching, but it can be overwhelming and inefficient to attempt to rate teachers on every single domain during each observation. Additionally, most rubrics do not take into account the complex, content-specific instructional moves that are known to help students master higher standards. Although many evaluators probably already triage the frameworks and look for signs of the instructional shifts associated with the Common Core and other new state standards, this focus is happening in spite of the system rather than as a result of it.

In evaluation systems that require multiple observations, some might be comprehensive but others could focus on specific elements of the teaching framework that represent the best opportunities for improvement—at least among teachers who have demonstrated baseline effectiveness—or target the biggest gaps in student learning. In **Colorado**, many districts have teachers do self-assessments first and then compare those with evaluator assessments during mid-year reviews. The process highlights discrepancies, areas of agreement, and topics to focus on during the next observation.²¹

Teachers and observers must agree on the purpose of each observation, but that requires acknowledging the tension between the developmental and accountability functions of evaluation. Teachers who are worried about their scores will be motivated to focus on areas of strength, and those who view evaluation as a mechanism for improving practice will want to concentrate on areas with the most potential for their own growth. Effective evaluation systems should be designed to balance these priorities. Comprehensive observations could be used as a check for baseline effectiveness, while scores of focused evaluations could be based on a teacher improving his or her ability to reflect on performance, to identify growth opportunities, and to progress over time. Such a structure might also allow teachers who already have demonstrated general effectiveness to concentrate on those areas with the greatest potential to improve practice without jeopardizing their performance ratings.

4 Allow for local discretion in accounting for student learning

Most teacher evaluation systems include a specific weight for student achievement data, even for teachers in non-tested grades and subjects, which is the overwhelming majority of teachers. For a variety of reasons, the policies related to accounting for student growth have been the most contentious in teacher evaluation debates. Over the past several years, tremendous energy and resources have gone into selecting and developing these measures, but those efforts have arguably created more tests, work, and bureaucratic compliance than they have increased student learning.²² And despite very different approaches, many states still have similar ranges in performance ratings.²³ It is important, however, not to let challenges associated with calculating student growth overshadow the need to account for student learning and to examine how these policy choices have played out.

Data collection and reporting, as well as communicating about complex statistical formulas, has been a challenge, especially as states have transitioned to new student assessments. Many states have already adjusted the weights or temporarily suspended the use of scores for high-stakes decisions as part of these transition periods, but in most

cases this was a reactive move made after significant pushback. States should proactively account for these testing transitions as part of broader implementation timelines both for new assessments and for teacher evaluation.

In addition to transition planning, states have struggled to determine how to calculate student growth, particularly for teachers in non-tested grades and subjects, but also for teachers of core subjects. Some states have set formula weights for both value-added measures (VAM) and other measures of student learning. Others have assigned a weight to a broader category of student learning data and left greater discretion for local decisions. Still other states rely on matrices or “body of evidence” requirements.²⁴ **Kentucky**, for example, provides a summative rating of teacher performance only after collecting three years of evidence in its Teacher Professional Growth and Evaluation System (TPGES). This body of evidence includes teacher self-reflection and professional growth planning, student achievement data, principal observations, student surveys, and other sources determined at the district level. The summative ratings are guided by “decision rules” determined at the state level to help evaluators make decisions informed both by professional judgment and multiple data sources.²⁵

Colorado requires 50 percent of teachers’ evaluations to be based on student academic growth data but created flexibility by leaving the choice of which specific assessments to be used up to local districts. Katy Anthes, Executive Director for Educator Effectiveness in the Colorado Department of Education, said, “What we talk about for districts is a measure that’s meaningful for them based on their context and their kids. We can get better about those metrics, but let’s start somewhere and use what you have trust in. For some, that might be a teacher-developed assessment. For others, that will be a vendor-developed test. Every district is actively weighing that tradeoff all the time.”²⁶

Given the implementation challenges and the need for continuous improvement, states might also consider establishing “innovation zones” that invite LEAs and charters to diverge from the statewide model, possibly using different weights or measures of student growth, as long as certain conditions are met. This approach could include additional training for teachers on developing classroom assessments and a research component on the costs/benefits of the alternate approaches. Although it is possible this would make evaluation results more difficult to compare across districts, it might foster experimentation and research and development of new approaches that have not yet been explored, while still allowing SEAs to track elements of more traditional student growth measures to determine their effect on teacher effectiveness ratings. SEAs may be better suited to provide the technical assistance and support LEAs need to select measures (i.e., offering different sample models in addition to flexibility for LEAs to experiment with different weights) than SEAs are to determining the exact weights and measures at the state level.

SEAs can contract with a university or technical assistance provider to regularly assess the quality, integrity, and usefulness of student assessments used to measure growth and publish reports for state boards, educators, and the public. Such dissemination of best practices, as well as the creation of communities of practice around local assessment development, could result in stronger evaluation measures while improving teachers’ knowledge of the underlying content expectations for students.

Moreover, partnerships with institutions of higher education could allow teacher preparation programs to identify needs and improve pre-service training, for example, regarding assessment literacy. This is an area in which many states have developed practices over the past several years.

“What we talk about for districts is a measure that’s meaningful for them based on their context and their kids. We can get better about those metrics, but let’s start somewhere and use what you have trust in.”

— KATY ANTHES, EXECUTIVE DIRECTOR FOR EDUCATOR EFFECTIVENESS, COLORADO DEPARTMENT OF EDUCATION

5 Respect the limitations of value-added data

Regardless of the weight or how it is calculated, value-added measures (VAM) have limitations, particularly when it comes to differentiating teacher performance within the broad “middle” range of performance. Therefore, value-added scores are most reliable and useful for informing decisions about teachers at the very high or very low end of effectiveness.²⁷

Most teachers fall in the middle range of value-added scores, so other measures should be decisive in their performance ratings. States can allow these teachers to use rolling averages of VAM estimates and/or allow teachers to use the upper range of the estimates to mitigate against measurement error.

To take advantage of the strength of VAM at the tails of the distribution, evaluation systems should flag cases where value-added data—with a high degree of confidence—identify teachers with lower-than-average effectiveness. In cases where these teachers receive favorable evaluations, system leaders should be required to review the results and collect additional information.

Additionally, SEAs can analyze statewide and district data to determine which schools and districts have the biggest mismatch between student achievement and observation scores and provide additional training to evaluators in those schools. In **Tennessee**, eight regional coaches work with evaluators in schools that show the greatest misalignment between observation scores and VAM results (e.g., are two or more levels apart on a five-level scale). Working with the coaches is voluntary, but principals in more than 100 schools across the state that demonstrated this misalignment have taken advantage of the service, which includes targeted support to principals around the observation rubric and scoring, and staff training. Alignment between observation scores and VAM results has improved in many of participating schools, and in some, student achievement has improved.²⁸ States could also use this approach to trigger a review for teachers with high VAM results and low observation scores.

When value-added results strongly indicate better-than-average teacher effectiveness with a high degree of confidence, SEAs should coordinate with LEA leaders to ensure that schools and districts recognize those teachers for their contributions and leverage that talent to spread effective practice in support of students and schools in the greatest need of effective teachers. States may also explore differentiating the frequency of required observations based on these results; for example, **Louisiana** requires at least two observations with feedback, but one observation may be waived for teachers who have earned a highly effective VAM rating.²⁹ Further, district leaders can propose policy solutions to support these efforts (i.e., leadership roles, funding for retention, incentives, or increased flexibility) and apply for federal funding to support this work under the new Teacher and School Leader Incentive Program authorized in ESSA.³⁰

6 Support locally developed measures while pursuing improvements in their creation and use

In an effort to generate measures of student learning for all teachers, many states encourage adoption of off-the-shelf tests, perhaps implicitly valuing reliability (i.e., measure is the same across contexts) over validity (i.e., assessing what you most want to assess, in terms of alignment to local curricular choices and college- and career-ready expectations).

States should consider how to better balance the types of assessments used, allowing teams of teachers, within and across schools, to develop progress measures or student learning objectives (SLOs) that align with local curricula and instruction. Such measures should, by design, support teacher reflection and development, a critical goal of new instructional improvement systems. Implementing SLOs has been a challenge for many districts and states and yet the work still presents an opportunity for practitioners to engage in rich discussions about goals for student learning and professional practice, as long as they have sufficient time and space to reflect, collaborate, and learn.³¹

Implementing SLOs presents an opportunity for practitioners to engage in rich discussions about goals for student learning and professional practice.

The **South Dakota** Department of Education partnered with the state’s teachers union, regional education services agencies (ESAs), and universities to host statewide trainings on developing SLOs and creating tasks and assessments to measure them. The partnership with the universities and ESAs also allowed teachers to receive graduate credit for their work on SLOs. The state Department of Education, ESAs, and union adjusted the trainings based on results and feedback and later integrated SLO training into their teacher preparation programs. According to officials with the Department, the most important factor was time. “We realized that the more time teachers get to practice [on SLOs], the better they’ll be,” said Carla Leingang, Administrator for Certification and Teacher Quality in the Division of Assessment and Accountability. “Teachers will acknowledge, after writing and implementing an SLO for the first time, they learn a lot about how to improve the process and how to be more reflective. Teachers need the opportunity to learn with no consequences so they can

grow in the process.” Leingang recommends giving teachers at least two years to learn about, develop, test, and improve SLOs before using the SLOs for accountability purposes.³²

When measures are developed locally, school leaders should prioritize students’ learning needs and hard-to-assess abilities and skills that are important markers of college- and career-readiness (e.g., research projects, oral presentations, and longer-term projects). States and districts can also acknowledge concerns about the reliability of locally developed assessments by weighting them differently within the evaluation system; in **District of Columbia Public Schools (DCPS)**, teachers in tested grades and subjects have 50 percent of their evaluations based on student growth data, while teachers in non-tested grades and subjects have that measure weighted at 15 percent, which recognizes the differences in the types of data available for each group of teachers.³³

If evaluators are allowed to use the data from SLOs as part of a body of evidence, one strategy for strengthening these practices over time (and continuing with a consistent weight on student achievement measures) would be to allow evaluators to assess a teacher’s expertise at developing or selecting appropriate measures of student learning and setting reasonably ambitious goals as part of the teacher’s evaluation. This approach is also a practical strategy to address the longstanding concern that educators need greater assessment literacy, but it requires that principals and other leaders also become more skilled at using assessments.

7 Make sure all important aspects of teaching performance are valued in evaluations

Teaching is complex. Evaluations should go beyond classroom observations and test scores to consider all aspects of teaching, including pedagogical and content knowledge, lesson planning, analysis of student data, collaboration, ability to develop students’ social and emotional well-being, and willingness to take intellectual risks and to use feedback to improve.

Teachers have incredibly important responsibilities to develop students’ social and emotional well-being, their intellectual confidence and willingness to take risks, and their character more generally. Recent research demonstrates that teachers’ effectiveness in instilling these qualities and behaviors in their students varies widely.³⁴ Evaluation processes need to consider these differences among teachers both to identify best practices and to demonstrate that these efforts are valued.

Based on a review of research-based social-emotional learning (SEL) programs and scholarship, the Center for Great Teachers and Leaders identified 10 instructional practices that support students’ social-emotional competencies and identified overlap between those 10 instructional practices and instructional practices as defined in three of the most common professional teaching frameworks: Classroom Assessment Scoring System (CLASS), Danielson’s Framework for Teaching, and Marzano’s Observational Protocol.³⁵ States and districts can use these crosswalks or create their own from the professional teaching frameworks used in their evaluation systems to understand how SEL-related aspects of teaching performance are considered in their current evaluation models.

Student and peer surveys are another way to account for other developmental aspects of teaching quality. The Measures of Effective Teaching Project (MET) study found that student surveys provide reliable information about dimensions of teacher quality not captured in observation rubrics or on student tests, such as caring and collaboration.³⁶ MET did not examine surveys with stakes for evaluation, and some teachers and scholars have resisted including these measures in formal evaluation processes as schools and districts struggle to integrate surveys in a way that acknowledges teachers’ concerns about using student feedback in high-stakes evaluations. However, the importance of the characteristics measured by the surveys and the increasing recognition that teachers vary significantly warrants further attention. As with other measures, evaluation systems should allow for professional judgment to guide decisions associated with student survey data, rather than a formulaic approach that treats surveys as a standalone measure.

In **Pittsburgh Public Schools**, 15 percent of teachers’ overall evaluation scores comes from results on the Tripod™ student surveys,³⁷ which are differentiated for early elementary, elementary, middle, and high school levels.

Teachers have incredibly important responsibilities to develop students’ social and emotional well-being, their intellectual confidence and willingness to take risks, and their character.

Pittsburgh teachers work with their principals to review their survey results, observation ratings, and student growth measures and identify areas of strength and potential growth. Teachers report learning more about their practice by seeing the classroom experience through their students' eyes. "As the national narrative pushes us to think more broadly about teachers' impact on student outcomes, skills, and competencies, surveys provide a richer, broader lens on teachers' classroom performance," shared Tara Tucci, Performance Management Systems Director for the district. The district is also using survey results from questions on school climate and student engagement to improve conditions for students.³⁸ **Georgia** also includes student survey results as part of the body of evidence evaluators can consider when determining final performance ratings for teachers, and state leaders report similar benefits in teacher reflection and student efficacy, even though the surveys do not account for a specific weight in the evaluation.³⁹

Teaching is a collaborative process in which teams of teachers, social workers, guidance counselors, librarians, school nurses, administrators, and others work together. Evaluation processes should consider teachers' ability to collaborate with peers to improve learning and school culture as well as other important competencies, such as willingness to improve and ability to diagnose students' learning needs. **DCPS** includes a measure of commitment to school community as 10 percent of teachers' evaluation ratings in tested and non-tested subjects and grades. This measure includes (1) support of local school initiatives, (2) support for special education and ELL programs, (3) high expectations for academics and behavior, (4) partnership with families, and (5) instructional collaboration.⁴⁰ Principals build off a district-provided rubric to specify what qualifies as success in meeting each of these components.

8 Engage teachers in improving teacher evaluation systems

State and district leaders need to know whether teachers consider the evaluation process meaningful, credible, and helpful in improving their performance. States should survey teachers regularly and examine the responses by school and by LEA to recognize and learn from instances where teachers express support for the system and to identify and investigate (e.g., through focus groups or visiting teams) cases where teachers report dissatisfaction.

Many states are generating similar or related data through workplace-conditions surveys but have not necessarily closed the loop by using that data effectively to improve teacher evaluation systems. In partnership with the New Teacher Center, **Kentucky** biannually administers the Teaching, Empowering, Leading, and Learning (TELL) survey to all teachers. Although TELL is not an official part of the teacher evaluation system, state leaders analyze the data and compare them with TPGES data to provide comprehensive feedback to schools. This includes, for example, cross-referencing feedback about teacher leadership opportunities from TELL data with the data from the leadership domain of the statewide teaching framework to help school administrators identify connections and make adjustments.⁴¹ Additionally, the Kentucky Teacher Effectiveness Steering Committee, which was initially charged with the design, development, and deployment of TPGES, continues to meet quarterly to provide feedback on implementation and monitoring of TPGES, including further integration and use of TELL data.⁴² "The more [the Kentucky Department of Education] is transparent with progress and struggles, the more support we have overall. TPGES is not a [Department] plan; it is a steering committee plan and people realize that," says

Amanda Ellis, Associate Commissioner for the Kentucky Office of Next Generation Learners.⁴³ Being transparent and communicating proactively about how data are being used also demonstrates an openness to feedback and an orientation to continuous improvement at the system level.

Some school systems, such as **DCPS**, also generate data at the observer level to ensure that professional learning and continuous improvement are emphasized in the process.⁴⁴ For instance, after an observation and feedback cycle, a teacher may be asked to indicate whether the observer identified areas of strength in their instruction from which to build, provided specific and actionable feedback, and helped to develop action steps.⁴⁵

State and district leaders need to know whether teachers consider the evaluation process meaningful, credible, and helpful in improving their performance.

Observers and evaluators also can be surveyed to gauge the health and functioning of the teacher evaluation system. For instance, system leaders should ask observers how much time they spend attending to bureaucratic obligations versus providing feedback and guidance to teachers, especially if deep conversations regarding instructional improvement are expected. Leadership time is a scarce resource and system leaders need to hear from front-line implementers and find new ways to support evaluators to continuously improve system design toward the goal of improving teaching and learning.

Adding a partner can add capacity and credibility to efforts to incorporate teacher and evaluator feedback. Universities, education advocacy organizations, teacher associations, and organizations such as Hope Street Group, the New Teacher Center, and the Southern Regional Education Board may be able to elicit perspectives and suggestions that would otherwise go unheard. In **Kentucky**, the state Department of Education, Center for Teacher Quality, Hope Street Group, and Kentucky Education Association have partnered to use TELL results to improve TPGES. This partnership includes follow-up surveys and focus groups with teachers in a select number of districts.⁴⁶ This type of partnership provides both a formal and continuous feedback system to engage teachers in improving implementation.

9 Develop measures for testing the integrity of evaluation system design and implementation

States should invest in and communicate about quality assurance and continuous improvement processes. For example, states can examine broad correlations between teacher ratings and value-added data across schools and districts and review other data not related to evaluation to corroborate evaluation findings, such as trends in no-stakes tests,

attendance and discipline data, and staff culture/workplace surveys, among others. Examining these measures can alert system leaders to possible gaming or cases of inordinate attention being paid to evaluation metrics without a corresponding focus on overall school improvement. States and districts should send visit teams to schools with divergent signals and learn from the practices in schools with strong results across the array of measures.

Where correlation between system-level value-added scores and performance ratings are weakest, leaders must determine the cause(s) of discrepancies between professional judgment and quantitative measures and collaborate with LEAs to strengthen training and exercise of judgment. **Massachusetts** has an Educator Evaluation Dashboard that combines data from multiple sources and allows for analysis of evaluation system implementation in each district, for all teacher performance levels, and across the four standards used in the state's system. Beginning in the 2016-17 school year, the state Department of Elementary and Secondary Education (DESE) will provide the dashboard to district leaders, allowing for comparisons of the relationship among variables, including educator evaluation ratings and student growth; educator evaluation ratings and prior student achievement; areas of consistency or discrepancy across schools within a district; comparisons in educator evaluation ratings of teachers of core subjects compared to teachers of non-core subjects; and comparisons in ratings across teachers from different demographic groups, including gender and age. The goal is to improve evaluation efforts at both the state and local level. DESE staff will provide support to districts on how to use the dashboard and acknowledge that some data may not be as useful to all districts, particularly those with small numbers of different types of teachers or other variables. The Department will

CONTINUOUS IMPROVEMENT CAN'T HAPPEN ALL THE TIME

SEAs should establish a clear framework for collecting feedback and other evidence and for developing a process and timeline to guide continuous improvement of teacher evaluation policies and practices (e.g., going from version 1.0 to 2.0). Initial decisions need time to run their course before impact and implications can be fully appreciated; if the evaluation system is in constant flux, it is difficult to assess the efficacy and consequences of various decisions. SEAs that develop a predictable, transparent process for assessing system effectiveness and adopting changes will create greater tolerance for implementation challenges and modifications.

continue to seek feedback on the usefulness of the data and the dashboard and continue to make improvements to the system and DESE services overall.⁴⁷

This type of ongoing analysis of teacher evaluation data is important; recent research has identified that teachers whose students have higher incoming achievement levels receive better classroom observation scores on average than those whose incoming students are at lower achievement levels. This trend indicates a potential bias in the observation system: When observers see a teacher leading a class with higher-ability students, they judge the teacher more favorably.⁴⁸

On a related note, states should have a research agenda or engage in research partnerships to determine the pros and cons of policy decisions.

10 Tell stories that go beyond performance ratings

Most news stories about teacher evaluations in the past several years prominently report the percentage of teachers assigned to each performance category, almost always emphasizing that very few teachers were deemed to be ineffective. This narrative improperly identifies the rating of teachers as the primary purpose of evaluations and implies that evaluation systems are too lenient.

Policymakers have long argued that the primary goal of evaluation systems is to improve teaching and learning by providing teachers with constructive feedback and identifying highly effective teaching practices. The data collected and reported should demonstrate this purpose with evidence of how teachers are improving their instructional practice over time.

Wisconsin is working to focus its teacher evaluations on meaningful supports for teachers, rather than on comparing and ranking educators based on overall evaluation scores. According to Katharine Rainey, Director of Educator Effectiveness for the Wisconsin Department of Public Instruction, “The only scores we think are important are tied to rubrics that tell you where you are, where you want to go, and how to get there by comparing specific language across the rubric levels.” She observed that when the process is focused on growth, it permeates all levels of the system, from the SEA to the students. In one district, students saw their teachers focused on data, cycles of inquiry, and goal-setting and adopted many of these habits for themselves.⁴⁹ By communicating that the process is about growth, vulnerability, and creating a safe space to continuously learn and improve, Wisconsin has helped change the conversation around the purposes of their evaluation system.

A positive narrative also helps parents and the public understand the fundamental value of teacher evaluation. Schools and districts should consider how they are communicating to families about the importance of teachers, teacher effectiveness, and evaluation efforts. When parents are informed, they are better empowered to support schools’ efforts, to advocate for more change or resources, and to hold leaders accountable when policies and processes do not lead to promised improvements in student learning.

Conclusion

The purpose of teacher evaluation is to accelerate professional growth and development that leads to instructional improvement and greater success for students, not to create anxiety and concerns about job security among educators.

licensure to rigorous standards of professional practice to tenure decisions, investments in professional growth and development, career ladders, and compensation. Teacher evaluation is a critical part of this work and should be effective and reliable, particularly given its prominence and cost. It is also important to make the set of policies around teacher effectiveness work together to set high expectations, attract and retain top talent, and ensure that all students benefit from highly effective teachers.

States have an important opportunity to clarify the purpose of teacher evaluation, to take stock of implementation experience, and to reconsider policy choices to improve operation and outcomes. If states do this well, they can bolster confidence in teacher evaluation and improve teacher learning and student achievement.

In some instances, however, new evaluation policy implementation has created a worst-of-both-worlds scenario: Exceedingly few teachers are identified as low-performing and at-risk of termination, but many teachers feel anxiety about this possibility, which undermines the learning orientation they should bring to the evaluation process. School systems need processes for identifying and potentially removing teachers who fall below expectations, but those methods must not leave all teachers feeling that their jobs are on the line every year.

Attempts to improve teacher evaluation processes should be undertaken, therefore, in the context of an overall policy framework for improving the educator workforce, from initial teacher preparation and

Acknowledgments

The Aspen Education & Society Program extends its deepest thanks to the state leaders, district administrators, practitioners, advocates, and thought partners who provided valuable insights and feedback. The authors alone are responsible for these recommendations and for any inaccuracies or errors.

- Katy Anthes, Executive Director, Educator Effectiveness, Courtney Cabrera, Principal Strategic Consultant, Educator Effectiveness, and Sed Keller, Assessment Implementation Specialist, Educator Effectiveness; Colorado Department of Education
- Andy Baxter, Vice President of Educator Effectiveness, Southern Regional Education Board
- Carl Blanchard, Director, Office of Evaluation, New Jersey Department of Education
- Mary Ann Blankenship, Executive Director, Kentucky Education Association
- Brad Clark, Director, Kentucky State Teacher Fellows Program, Hope Street Group
- Rachel Curtis, Founder, Human Capital Strategies for Urban Schools
- Hannah Dietsch, Assistant Superintendent, Office of Talent, Louisiana Department of Education
- Amanda Ellis, Associate Commissioner, Office of Next Generation Learners, Kentucky Department of Education
- Paul Fleming, Assistant Commissioner, Division of Teachers and Leaders, and Courtney Seiler, former Deputy Director of Evaluation; Tennessee Department of Education
- Suraj Gopal, Teacher, Hudson High School of Learning Technologies, New York City Department of Education
- Sara Heyburn, Executive Director, Tennessee State Board of Education
- Michelle Hudacsko, Deputy Chief of Instructional Evaluation, District of Columbia Public Schools
- Misti Kemmer, Teacher, Russell Elementary School, Los Angeles Unified School District
- Carla Leingang, Administrator for Certification and Teacher Quality in the Division of Assessment and Accountability, and Matt Gill, Educator Effectiveness Specialist; South Dakota Department of Education
- Patrick McGuinn, Associate Professor of Political Science, Drew University
- Chris Minnich, Executive Director, Mary-Dean Barringer, Strategic Initiative Director, and Holly Boffy, Program Director, Teacher and Leader Development; Council of Chief State School Officers
- Kari Patrick, Teacher, Shelby County High School, Shelby County Public Schools (Kentucky)
- Heather Peske, Associate Commissioner of Educator Effectiveness, and Ron Noble, Manager, Educator Evaluation; Massachusetts Department of Elementary and Secondary Education
- Stephanie Plachy, Teacher, Evergreen Middle School for Urban Exploration, New York City Department of Education
- Katharine Rainey, Director, Educator Effectiveness, Wisconsin Department of Public Instruction
- Christopher Ruzzkowski, Chief Officer, Teacher and Leader Effectiveness Unit, Laura Schneider, Special Assistant of Educator Effectiveness, and Shanna Ricketts, Data Strategist; Delaware Department of Education
- Sara Sanchez, Managing Director of Policy, Educators 4 Excellence – Chicago
- Mary Ann Snider, Chief of Educator Excellence and Instructional Effectiveness, and Lauren Matlach, Education Specialist, Office of Educator Quality and Certification; Rhode Island Department of Education
- Tara Tucci, Performance Management and Systems Director, Pittsburgh Public Schools
- Joanne Weiss, Weiss Associates LLC
- Jess Wood, Senior Policy Advisor, EducationCounsel

From the Aspen Education & Society Program team, we acknowledge Ross Wiener, Danielle Gonzales, Marisa Goldstein, and Lee Kappes for their contributions to the report.

Richard Colvin reviewed an earlier draft of this report and provided valuable feedback. Next Chapter Communications designed the report.

The Aspen Education & Society Program is grateful to the Bill & Melinda Gates Foundation and the Joyce Foundation for their financial support of our work.

Any opinions, recommendations, or conclusions expressed in this material are those of the authors and do not necessarily reflect the views of these foundations.

Endnotes

1. For more information, see Lauren Camera, “Waiver States Can Seek a Year’s Delay in Using Tests for Teacher Evaluation,” *Education Week*, August 21, 2014, available from http://blogs.edweek.org/edweek/campaign-k-12/2014/08/ed_announcement.html and Kate Taylor, “Cuomo, in Shift, Is Said to Back Reducing Test Scores’ Role in Teacher Reviews,” *New York Times*, November 25, 2015, available from http://www.nytimes.com/2015/11/26/nyregion/cuomo-in-shift-is-said-to-back-reducing-test-scores-role-in-teacher-reviews.html?_r=0.
2. Stephen Sawchuk, “ESSA Loosens Reins on Teacher Evaluations, Qualifications,” *Education Week*, January 5, 2016, available from <http://www.edweek.org/ew/articles/2016/01/06/essa-loosens-reins-on-teacher-evaluations-qualifications.html>.
3. For additional guidance on how SEAs can decide what responsibilities to take on and what to delegate to LEAs, see “Roles and Responsibilities of the State Education Agency,” Aspen Institute Education & Society Program, December 2015, available from <http://www.aspendrl.org/portal/browse/DocumentDetail?documentId=2846&download&admin=2846|768802984>.
4. Mona Mourshed, Chinezi Chijoke, and Michael Barber, “How the World’s Most Improved School Systems Keep Getting Better,” McKinsey & Company, November 2010, available from http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/social%20sector/pdfs/how-the-worlds-most-improved-school-systems-keep-getting-better_download-version_final.ashx. The six strategies described in the report include strengthening the technical skills of new and experienced teachers and principals; assessing students; using data to inform improvement; revising standards and curricula; introducing differentiated compensation models; and articulating the vision for improving the education system in laws and policies.
5. Melinda Adnot et al., “Teacher Turnover, Teacher Quality, and Student Achievement in DCPS,” National Bureau of Economic Research Working Paper No. 21922, January 2016, available from <http://www.nber.org/papers/w21922>.
6. Emma Brown, “D.C. Public Schools, Closely Watched for Its Reform Efforts, Is Overhauling Teacher Evaluation and Training,” *Washington Post*, February 10, 2016, available from https://www.washingtonpost.com/local/education/dc-public-schools-to-overhaul-teacher-evaluation-and-training/2016/02/10/bdb9ed2a-cf41-11e5-b2bc-988409ee911b_story.html.
7. For more information, see the Council of Chief State School Officers, “Principles for Teacher Support & Evaluation Systems,” March 1, 2016, available from <http://www.ccsso.org/Documents/2016/Principles%20for%20Teacher%20Support%20and%20Evaluation%20Systems.pdf>.
8. Kathryn M. Doherty and Sandi Jacobs, “State of the States 2015: Evaluating Teaching, Leading and Learning,” National Council on Teacher Quality, November 2015, p. 96, available from <http://www.nctq.org/dmsView/StateofStates2015>.
9. Lauren Sartain et al., “Rethinking Teacher Evaluation in Chicago: Lessons Learned from Classroom Observations, Principal-Teacher Conferences, and District Implementation,” Consortium on Chicago School Research, November 2011, available from <https://consortium.uchicago.edu/sites/default/files/publications/Teacher%20Eval%20Report%20FINAL.pdf>.
10. For more information on NIET’s engagement with Tennessee on the TAP Teaching Standards, see <http://www.niet.org/best-practices-center/bpc-partners/>.
11. Heather Peske (Associate Commissioner of Educator Effectiveness, Massachusetts Department of Elementary and Secondary Education) and Ron Noble (Manager of Educator Evaluation, Massachusetts Department of Elementary and Secondary Education), interview with authors, January 20, 2016. The Massachusetts Performance Assessment for Leaders, a requirement for principal licensure, includes two of four performance tasks related to providing feedback to teachers and to building a professional learning culture. For more information, see www.ma-pal.com.
12. For more information see “Quick Reference Guide: Educator Plans,” Massachusetts Department of Elementary and Secondary Education, September 2014, available from <http://www.doe.mass.edu/edeval/resources/QRG-EducatorPlans.pdf>.
13. Heather Peske and Ron Noble, interview with authors.
14. For more information, see Harvard Graduate School of Education, “A User’s Guide to Peer Assistance and Review (PAR),” available from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf.
15. John P. Papay et al., “Learning Job Skills from Colleagues at Work: Evidence from a Field Experiment Using Teacher Performance Data,” National Bureau of Economic Research Working Paper No. 21986, February 2016, available from <http://www.nber.org/papers/w21986>.
16. TNTP, “The Irreplaceables: Understanding the Real Retention Crisis in America’s Urban Schools,” July 30, 2012, available from http://tntp.org/assets/documents/TNTP_Irreplaceables_2012.pdf.
17. Heather Peske and Ron Noble, interview with authors. For more information on professional growth plans for educators in Massachusetts, see: <http://www.doe.mass.edu/edeval/resources/QRG-EducatorPlans.pdf>.
18. Rachel Curtis, “Achievement First: Developing a Teacher Performance Management System that Recognizes Excellence,” The Aspen Institute Education & Society Program, March 2011, available from <http://www.aspendrl.org/portal/browse/DocumentDetail?documentId=1510&download>.
19. Ben Jensen et al., “Beyond PD: Teacher Professional Learning in High-Performing Systems,” Learning First, January 2016, available from <http://www.ncee.org/wp-content/uploads/2015/08/BeyondPDWeb.pdf>.
20. For more information on how states, districts, and schools can create teacher leadership roles that increase system capacity and are well-defined and resourced, see “Leading from the Front of the Classroom: A Roadmap for Teacher Leadership that Works,” The Aspen Institute Education & Society Program and Leading Educators, November 2014, available from <http://www.aspendrl.org/portal/browse/DocumentDetail?documentId=2402&download>.
21. Katy Anthes (Executive Director for Educator Effectiveness, Colorado Department of Education), interview with authors, January 22, 2016.
22. Ray Hart et al., “Student Testing in America’s Great City Schools: An Inventory and Preliminary Analysis,” Council of the Great City Schools, October 2015, available from <http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Testing%20Report.pdf>.

23. Matthew A. Kraft and Allison F. Gilmour, "Revisiting the Widget Effect: Teacher Evaluation Reforms and the Distribution of Teacher Effectiveness," Brown University Working Paper, February 2016, available from http://scholar.harvard.edu/files/mkraft/files/kraft_gilmour_2016_revisiting_the_widget_effect_wp.pdf?m=1456772152.
24. In 2015, 43 states required objective measures of student growth as a factor in teacher evaluations. For more information on how states weight student growth in principal and teacher evaluation and how these growth measures are defined, see Kathryn M. Doherty and Sandi Jacobs, "State of the States 2015: Evaluating Teaching, Leading and Learning," National Council on Teacher Quality, November 2015, p. 96, available from <http://www.nctq.org/dmsView/StateofStates2015>.
25. "Summative Evaluation Process," Kentucky Department of Education, January 26, 2016, available from <http://education.ky.gov/teachers/PGES/TPGES/Pages/TPGES-Summative-Evaluation-Process.aspx>.
26. Katy Anthes, interview with authors.
27. Andrea Lash et al., "Analysis of the Stability of Teacher-level Growth Scores from the Student Growth Percentile Model," National Center for Education Evaluation and Regional Assistance and Regional Educational Laboratory at West Ed, January 2016, available from https://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2016104.pdf.
28. Paul Fleming (Assistant Commissioner in the Division of Teachers and Leaders, Tennessee Department of Education), interview with authors, January 15, 2016.
29. Sandi Jacobs et al., "2015 State Teacher Policy Yearbook: Louisiana," National Council on Teacher Quality, December 2015, available from http://www.nctq.org/dmsView/Louisiana_2015_NCTQ_State_Teacher_Policy_Yearbook.
30. The Teacher and School Leader Incentive Fund authorized by the Every Student Succeeds Act (ESSA) replaces the Teacher Incentive Fund (TIF) that existed under No Child Left Behind. For more information on the changes to TIF under ESSA, see the Council of Chief State School Officers, "Comparison of Select Elements of ESEA," December 8, 2015, available from <http://www.ccsso.org/Documents/2016/ESSA/CCSSOComparisonofSelectElementsofESEA12142015.pdf>.
31. For more information on SLOs, see Lisa Lachlan-Haché, Ellen Cushing, and Lauren Bivona, "Student Learning Objectives: Benefits, Challenges, and Solutions," American Institutes for Research, November 2012, available from http://educatoralent.org/inc/docs/SLOs_Benefits_Challenges_Solutions.pdf.
32. Carla Leingang (Administrator for Certification and Teacher Quality in the Division of Assessment and Accountability, South Dakota Department of Education), interview with authors, January 20, 2016.
33. Erika Hall et al., "State Practices Related to the Use of Student Achievement Measures in the Evaluation of Teachers in Non-Tested Subjects and Grades," National Center for the Improvement of Educational Assessment, August 26, 2014, available from http://www.nciea.org/publication_PDFs/Gates%20NTGS_Hall%20082614.pdf.
34. Recent research shows that teachers have a differential impact on students' attendance, discipline, and grade promotion, factors that predict long-term success of students as much as growth on test scores. For more information, see C. Kirabo Jackson, "Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina," Northwestern University and National Bureau of Economic Research, July 2014, available from <http://www.nber.org/papers/w18624.pdf>.
35. Dr. Nicholas Yoder, "Teaching the Whole Child: Instructional Practices That Support Social-Emotional Learning in Three Teacher Evaluation Frameworks," Center for Great Teachers and Leaders, January 2014, available from <http://www.gtlcenter.org/sites/default/files/TeachingtheWholeChild.pdf>.
36. "Feedback for Better Teaching," Bill & Melinda Gates Foundation, January 2013, available from <http://files.eric.ed.gov/fulltext/ED540956.pdf>.
37. More information on the Tripod frameworks and survey modules is available from <http://tripoded.com/>.
38. Tara Tucci (Performance Management and Systems Director, Pittsburgh Public Schools), interview with authors, February 11, 2016.
39. Cynthia Saxon (Associate Superintendent of Teacher and Leader Effectiveness, Georgia Department of Education), interview with authors, February 23, 2016.
40. Michelle Hudacsko (Deputy Chief of Instructional Evaluation, District of Columbia Public Schools), interview with authors, January 22, 2016.
41. Amanda Ellis, interview with authors.
42. Mary Ann Blankenship (Executive Director, Kentucky Education Association), interview with authors, January 15, 2016.
43. Amanda Ellis, interview with authors.
44. Michelle Hudacsko, interview with authors.
45. Ross Wiener and Kasia Lundy, "Evaluating Evaluations: Using Teacher Surveys to Strengthen Implementation," The Aspen Institute Education & Society Program and The Parthenon Group, April 22, 2013, available from <http://www.aspendri.org/portal/browse/DocumentDetail?documentId=1597&download>.
46. Brad Clark (Director, Kentucky State Teacher Fellow Program, Hope Street Group), interview with authors, January 25, 2016; Amanda Ellis, interview with authors.
47. Heather Peske and Ron Noble, interview with authors.
48. Matthew P. Steinberg and Rachel Garrett, "Classroom Composition and Measured Teacher Performance: What Do Teacher Observation Scores Really Measure?" Educational Evaluation and Policy Analysis (January 2016), available from <http://static.politico.com/58/5f/fl4b2b144846a9b3365b8f2b0897/study-of-classroom-observations-of-teachers.pdf>.
49. Katharine Rainey (Director of Educator Effectiveness, Wisconsin Department of Public Instruction), interview with authors, January 21, 2016.



THE ASPEN INSTITUTE

EDUCATION & SOCIETY PROGRAM