KAYLAN CONNALLY AND MELISSA TOOLEY

BEYOND RATINGS

Re-envisioning State Teacher Evaluation Systems as Tools for Professional Growth

MARCH 2016

About the Authors

Kaylan Connally is a policy analyst with New America's Education Policy Program. She provides research and analysis on educator quality policies along with other PreK-12 issues. Connally holds an Ed.M. in education policy and management from the Harvard Graduate School of Education and a B.A. in philosophy from New York University.

Melissa Tooley is a senior policy analyst with New America's Education Policy Program. She provides research and analysis on K-12 policies and practices that impact teaching quality and school leadership. Tooley completed her master's in public policy at the University of Chicago and holds a bachelor's degree in psychology from Drew University.

Acknowledgements

About New America

New America is committed to renewing American politics, prosperity, and purpose in the Digital Age. We generate big ideas, bridge the gap between technology and policy, and curate broad public conversation. We combine the best of a policy research institute, technology laboratory, public forum, media platform, and a venture capital fund for ideas. We are a distinctive community of thinkers, writers, researchers, technologists, and community activists who believe deeply in the possibility of American renewal.

Find out more at **newamerica.org/our-story**.

About the Education Policy Program

New America's Education Policy Program uses original research and policy analysis to solve the nation's critical education problems, serving as a trusted source of objective analysis and innovative ideas for policymakers, educators, and the public at large. We combine a steadfast concern for low-income and historically disadvantaged people with a belief that better information about education can vastly improve both the policies that govern educational institutions and the quality of learning itself. Our work encompasses the full range of educational opportunities, from early learning to primary and secondary education, college, and the workforce. We are deeply engaged in ongoing developments in educational technology at all levels of child and adult development. We believe new organizational models have potential to achieve breakthrough levels of performance on behalf of students. And we believe that all providers of education must be held constructively accountable for the quality of their work.

Find out more at **newamerica.org/education-policy**.

Contents

INTRODUCTION AND SUMMARY

The spotlight on the American public school teacher has never been brighter. To ensure that all students have access to quality teaching, most states have adopted new, more rigorous teacher evaluation systems over the past five years based on multiple measures of teacher performance, such as evidence of student learning and observations of teacher practice. At the same time, most states have also adopted more rigorous academic standards in order to better prepare students for college, career, and civic life. These two efforts fit together: students will meet the new academic expectations only insofar as their teachers can successfully deliver on them in the classroom.

Given the critical need to improve student learning and the outsized influence that teachers have in this area, both efforts are imperative.[1] But evaluation systems put teachers under heightened scrutiny at the same time that they are revisiting how and what they teach. Carrying out new college- and career-ready standards will require all teachers—irrespective of content area—to make planning and pedagogical shifts, with an even greater emphasis on helping students develop their problem-solving and analytic skills through effective questioning and discussion techniques.[2]

To help teachers meet this challenge, states, local education agencies (LEAs), and schools must work together to ensure that teachers receive frequent, targeted feedback and suggestions for how to grow their practice.[3] Fortunately, teacher evaluation systems hold much potential for delivering the kind of constructive feedback and aligned learning opportunities that can promote teacher improvement.

However, to date, most of the public narrative and teacher pushback about evaluation has centered on its use for high-stakes personnel decisions such as pay, promotion, and dismissal. Given this, some might be puzzled, or even bristle, at the suggestion that data from evaluation systems could be used to drive instructional improvement. But the federal policies that encouraged states to adopt new evaluation systems always intended for these systems to raise the quality of teaching through not just accountability *but also support*: to recognize and reward the very top teachers, dismiss those at the very bottom, *and* promote improvements for those in between. And many states followed through with policies that reflect that intent: in a 2015 review, the National Council on Teacher Quality (NCTQ) found that 30 states and the District of Columbia require teachers' evaluations to inform their professional development.[4]

If this is the case, why is teacher development still missing from the public narrative on—and most teachers' daily experience of—evaluation? It is because while policy can provide a necessary foundation, it is insufficient without follow-through by states and LEAs. This paper examines the policies and practices that those 31 states* first identified by NCTQ have in place to assist LEAs in making the connection between teacher evaluation and support.

By conducting in-depth interviews with state education agency leaders in nearly all of those states, we found that, to date, the majority have not yet helped LEAs make good on these policies. For the most part, states have prioritized getting evaluation systems up and running and are only beginning to think about using them to promote ongoing teacher learning and growth. This is not altogether surprising; for many states, setting up the basic components of a teacher evaluation system has been a heavy lift in and of itself, particularly alongside instituting new academic standards and the associated student assessments.

Despite the fact that very few teachers have been dismissed under new evaluation systems, these initial evaluation efforts have bred distrust among some educators.[5] And the rise of new academic standards and accompanying assessments have made many teachers more anxious about the use of evaluation for accountability. While most educators believe in the goals of new college- and career-ready standards, they do not want to be set up to fail. For educators to succeed, they want and need targeted support.[6]

Fortunately, a few states that are further along in implementing their evaluation systems are beginning to use the information gleaned to go beyond performance ratings to include a focus on teacher development. While these efforts have taken different shapes depending on the state context, **all are focused on some combination of communication, support, and monitoring** of evaluation as a tool to support teacher growth. This paper highlights a representative sample of what we learned from states and digs deeper into the work of four—Colorado, Delaware, Louisiana, and Tennessee—to share promising practices and lessons learned.

But even for the states highlighted, more work can be done to meaningfully connect evaluation with support. Despite LEAs and schools holding primary responsibility for teachers' ongoing improvement, states have a substantial role to play in helping ensure that evaluation systems shift to support teacher growth in meaningful ways. States should strive to put policies in place to ensure teachers receive frequent, accurate feedback and to communicate to educators that a key purpose of evaluation is to support teacher growth, in addition to accountability. Then, states must ensure that LEA and school leaders follow through, by providing them with training, tools, and resources to inform teacher development, and by monitoring local efforts and outcomes.

We recognize that states' spheres of influence will vary depending on their context and that they will not be able to take all of the necessary steps at once. The work will require thoughtful prioritization, and consideration of how it can be integrated with other education initiatives, in the face of funding, capacity, and logistical constraints. Some states may feel that the recent passage of the Every Student Succeeds Act (ESSA), which reauthorized the federal Elementary and Secondary Education Act, will make this work even harder, as it ended the strongest federal incentives to develop and implement more rigorous educator evaluation systems.

> **States should embrace this responsibility and use it as an opportunity to retool their teacher evaluation systems to include a clear focus on professional growth, in addition to accountability.**

However, the law provides funding for states to invest in systems that "provide useful and timely feedback and... inform decision making about professional development [and] improvement strategies." In fact, by enacting ESSA, the federal government has put the obligation for ensuring teaching quality and student success squarely on states' shoulders, along with the expectation that states will fulfill it. States should embrace this responsibility and use it as an opportunity to retool their teacher evaluation systems to include a clear focus on professional growth, in addition to accountability. By moving forward in this area, states can help ensure that all teachers—and their students—succeed.

 * For the remainder of the paper, we use "31 states" when referencing the 30 states and the District of Columbia's Office of the State Superintendent identified by NCTQ as having policies requiring that teacher evaluation inform professional development. Although we attempted to contact and interview individuals at all 31 state education agencies, five states were non-responders (see Interviews Conducted, page 34).

WHY EVALUATION SYSTEMS MUST EMBRACE A FOCUS ON SUPPORT

The federal initiatives that helped incentivize new multiple-measure evaluation systems intended for them to have a dual purpose of both accountability and support (see: "The Rise of New Teacher Evaluation Systems and How Development Got Largely Overlooked" on page 6). But to date, states have primarily focused on the former. The majority of states concentrated their efforts on designing and operationalizing these new systems—for example, developing measures of student growth—in order to rate teachers' performance and identify the proportion of teachers who fall in each rating category. (See: "What Are Student Growth Measures?" on page 5.) Some states are starting to use these ratings to identify top teachers for recognition and rewards and bottom teachers for remediation and dismissal.[8] For instance, Tennessee is using its evaluation system to identify high-performing teachers for participation in a teacher leadership network, and local education agencies have also begun to dismiss a small number of persistently low-performers.[9]

But in most states the results from new multiple-measure teacher evaluation systems show that the vast majority of teachers are being rated in the top rating categories, leaving little differentiation on which to base personnel decisions like these.[10] In particular, very few teachers have been identified and dismissed for poor performance under new evaluation systems despite fear and anxiety among teachers that they would be.[11]

Teachers continue to be rated highly under new teacher evaluation systems for several reasons, including system design and the need for better observer training. Beyond these technical factors, there are interpersonal dynamics at play as well. School principals and other observers** charged with evaluating teacher practice often face disincentives to providing accurate ratings, including the desire to maintain strong staff morale.[12] But, even if all these factors improved tomorrow and evaluation systems captured a more nuanced picture of teacher performance, most of the 3.5 million teachers in the U.S. would likely still fall in the middle of the performance distribution, as neither superstars nor laggards (see Figure 1).[13] While it is unreasonable to expect all teachers to become superstars, states, LEAs, and schools can provide high-quality targeted supports to help teachers become better at their craft and help shift the entire distribution of effective teaching forward.

**We use the term "observer" rather than "evaluator" as the classroom observation component of evaluation systems is more likely to provide rich data and feedback to support teacher improvement than the summative rating alone. Most often the observer tasked with providing a final evaluation score is the school principal, although some systems allow for multiple evaluators or observers as discussed in the following section.

What Are Student Growth Measures?

This paper focuses primarily on the observation component of new multiple-measure teacher evaluation systems, as observation feedback has high potential for supporting teacher improvement and generally applies to all teachers. However, public debate surrounding evaluations has primarily focused on another system component that a majority of states now require: student growth measures. Though these measures take on a variety of names and forms, they generally refer to students' academic progress between two points in time, typically from one year to the next.[i]

In most states, for teachers in grades and subjects with annual standardized assessments, the "student growth" measure is based largely or wholly on their students' progress on these assessments. The models used to calculate growth range in complexity and their ability to isolate an individual teacher's impact—from simple comparisons and descriptive analyses to more sophisticated value-added models that attempt to control for other outside factors.[ii]

For teachers in "untested" grades and subjects, 22 states allow or require "student learning objectives" (SLOs) to be used to measure student growth.[iii] SLOs are specific student learning goals accompanied by measures to track progress toward that goal.[iv] Typically based on assessments other than annual statewide tests, they are often set by a teacher at the beginning of the school year and dependent on the school leader's approval.[v] These may be used for teachers in "tested" grades and subjects as well.

i Jim Hull, "Measuring Student Growth: A Guide to Informed Decision Making," (Alexandria, VA: Center for Public Education, November 2007); Battelle for Kids, "Student Growth Measures in Ohio."

ii Mark Ehlert, Cory Koedel, Eric Parsons, and Michael Podgursky, *Selecting Growth Measures for School and Teacher Evaluations: Should Proportionality Matter?* (Washington, DC: National Center for Analysis of Longitudinal Data in Education Research, May 2013).
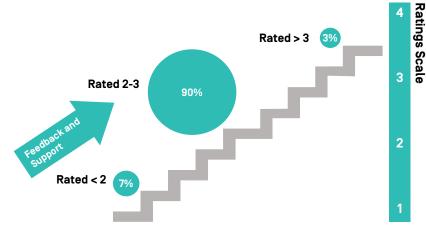
iii Kathryn M. Doherty and Sandi Jacobs, *State of the States 2015: Evaluating, Teaching, Leading and Learning* (Washington, DC: National Council on Teacher Quality, November 2015).

iv Reform Support Network, "Targeting Growth: Using Student Learning Objectives as a Measure of Educator Effectiveness."

v Morgaen L. Donaldson and John P. Papay, "Teacher Accountability for Evaluation and Research," chap. 11 in *Handbook of Research in Education Finance and Policy*, edited by Helen F. Ladd and Margaret E. Goertz (New York: Routledge, 2014); Lisa Lachlan-Haché, Ellen Cushing, and Lauren Bivona, *Student Learning Objectives as Measures of Educator Effectiveness: The Basics* (Washington, DC: American Institutes for Research, 2012).

Figure 1: Most Teachers are Neither Superstars nor Laggards

Summative Observation Ratings of MET Project Teachers

Rated > 3 — 3%
Rated 2-3 — 90%
Rated < 2 — 7%

Feedback and Support

Ratings Scale: 4, 3, 2, 1

Over three years, the Measures of Effective Teaching (MET) project studied how best to identify and support great teaching practice using data from multiple-measure evaluation systems, including classroom observations, student achievement gains, and student surveys. The study found that, even when observed by independent, trained, and certified observers, the vast majority of teachers still fell in the middle of the performance distribution.

Note: Adapted from Education Counsel's report, "The Quality Framework," April 2014.

The Rise of New Teacher Evaluation Systems and How Development Got Largely Overlooked

As highlighted in TNTP 's seminal report, *The Widget Effect*, prior to 2009, most states and LEAs lacked a robust method for measuring teacher performance. Principals typically performed cursory observations of teachers' classroom practice once a year, using compliance-oriented checklists, while in some places, tenured teachers went unobserved. Frequently the areas included on these checklists had little to do with teachers' ability to help students learn—for example, the orderliness of teachers' bulletin boards rather than observable teacher behaviors—and no other measures of teachers' impact on students' learning were included in the evaluation.[i] Principals were then required to provide teachers with an overall rating on what was often a binary scale (e.g., "satisfactory" vs. "unsatisfactory").[ii] Not surprisingly, these older state teacher evaluation systems tended to rate nearly all teachers as satisfactory.[iii] Because the ratings from these systems masked real differences in educator effectiveness, they were rarely used to inform personnel decisions such as tenure, dismissal, or promotion. And because the systems provided no mechanisms for concrete, evidence-based feedback on teacher practice, they generally did not inform professional development for teachers.

Spurred in part by TNTP's findings, the federal government launched two major initiatives to improve state teacher evaluation policy.[iv] First, the American Recovery and Reinvestment Act of 2009 funded Race to the Top (RTT), which provided competitive grants for states to voluntarily adopt policies and practices aligned with the Department of Education's (ED) priorities.[v] States could earn points for adopting policies and practices in specific categories, with the "supporting great teachers and leaders" category—encompassing everything from teacher preparation to evaluation and support—representing over a quarter of all points allocated.[vi] And a majority of points within this category could be earned by efforts to improve teacher and principal effectiveness through performance evaluation systems. The RTT performance evaluation criteria included designing systems that were: fair and rigorous; based on multiple rating categories that took into account multiple measures, including student growth as a "significant factor;" conducted annually to provide "timely and constructive feedback;" and used to inform staffing decisions including "*staff development*,

compensation, promotion, tenure, certification, and removal of ineffective teachers."[vii] [our emphasis] Since using evaluations to inform staffing decisions was worth the most points in the performance evaluation sub-category, all states that won a RTT grant had to have relatively strong applications in this area. As a result, most states' applications included plans for using evaluation to inform professional development. However, while ED took action when states failed to follow through on some elements of their teacher and leader effectiveness plans—such as withholding funds when Georgia failed to implement systemic merit pay tied to evaluation results—ED did not intervene in states that failed to follow through on plans to connect evaluation to professional development.[ix]

Second, in 2011, the Department of Education announced that it would grant states flexibility in implementing some of the requirements of the Elementary and Secondary Education Act (ESEA) via waivers from the law. To continue its teacher effectiveness push, ED made the adoption of multi-measure teacher evaluation systems, which include student growth as a "significant factor" in determining a teacher's final evaluation rating, one condition of the waivers.[x]

> **While ED took action when states failed to follow through on some elements of their teacher and leader effectiveness plans … ED did not intervene in states that failed to connect evaluation to professional development.**

While debates continue about whether U.S. Secretary of Education Arne Duncan was within his authority to put stipulations on states to earn an ESEA waiver, the efforts clearly had an impact on state policy. Just six years after *The Widget Effect*, 43 states have developed more rigorous evaluation systems that incorporate a measure of student learning growth and also include more detailed evidence-based classroom observations of teachers' practice.[xi] While about half of these

states had revised their evaluation systems before the Department announced its waivers, the remainder did so in large part, if not entirely, in response to the waivers.[xii]

However, many of these new systems are still not providing teachers with richer, more frequent feedback on their practice than they were before or differentiating teacher performance to inform their development. While the Department intended for states' waiver plans to catalyze better evaluation _and_ improvement activities, it only provided specific guidance on the former. ED required states to develop new teacher evaluation systems that included at least three levels of performance ratings, incorporated "multiple valid measures in determining performance levels, including as a significant factor data on student growth," and used evaluation ratings to inform personnel decisions by a certain date.[xiii] Less clear were the requirements stipulating evaluation systems "be used for continual improvement of instruction" and "provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development." And while—at least initially—ED denied some states waivers for not having the legal authority to ensure local implementation or for not implementing certain aspects of evaluation systems, such as student growth, within the specified timeline, no state was denied a waiver because it had not sufficiently planned for how to use these new systems to drive educator improvement.

i Linda Darling-Hammond, *Creating a Comprehensive System for Evaluating and Supporting Effective Teaching* (Palo Alto, CA: Stanford Center for Opportunity Policy in Education, 2012).

ii Daniel Weisberg, Susan Sexton, Jennifer Mulhern, and David Keeling, *The Widget Effect: The National Failure to Acknowledge and Act on the Differences in Teacher Effectiveness* (New York: The New Teacher Project, 2009).

iii Ibid.

iv Chad Alderman and Carolyn Chuong, *Teacher Evaluations in an Era of Rapid Change: From "Unsatisfactory" to "Needs Improvement"* (Washington, DC: Bellwether Education Partners, August 2014).

v National Center for Education Evaluation and Regional Assistance, "State Requirements for Teacher Evaluation Policies Promoted by Race to the Top," April 2014.

vi U.S. Department of Education, "Appendix B: Scoring Rubric," Race to the Top, January 2010.

vii National Center for Education Evaluation and Regional Assistance, "State Requirements for Teacher Evaluation Policies Promoted by Race to the Top," April 2014.

viii Alyson Klein, "Georgia Battles to Beat Race to Top Head Winds," *Education Week*, June 3, 2014; Charles Doolittle, e-mail to Kaylan Connally, February 2, 2016.

ix Kaitlin Pennington, *ESEA Waivers and Teacher-Evaluation Plans* (Washington, DC: Center for American Progress, May 2014).

x Kathryn M. Doherty and Sandi Jacobs, *2015 State of the States: Evaluating Teaching, Leading, and Learning* (Washington, DC: National Council on Teacher Quality, November 2015).

xi Ibid.

xii U.S. Department of Education, "Laws and Guidance: ESEA Flexibility."

xiii Jim Hull, "Measuring Student Growth: A Guide to Informed Decision Making," Center for Public Education, Nov. 9, 2007.

Despite this reality, the dominant narrative around evaluation has been about whether teachers are being fairly rated as "ineffective" or "in need of improvement" rather than the ongoing development of *all* teachers.[14] Unfortunately, current development policies tend to reinforce rather than combat this narrative: states and LEAs often only require "improvement plans" for their lowest-performing teachers, and schools often only intervene with support for struggling teachers.[15] As a result, the term "improvement" has developed a negative connotation within the context of teacher evaluation, undermining its potential as a vehicle for continuous teacher learning and growth.

While new teacher evaluation systems are still relatively young, some evidence exists that investing in better systems for both accountability *and* development could help teachers (and their students) improve, even more experienced ones. For instance, a study conducted by economists Jim Wyckoff and Thomas Dee found that the evaluation system in the District of Columbia Public Schools caused the voluntary attrition of low-performing teachers and marked improvements by those low-performers who remained teaching.[16] What is more, the teachers who replaced those who voluntarily left improved student achievement by meaningfully larger, statistically-significant amounts.[17] While the system included multiple observations by trained observers, the authors identified the consequences in place for poor performance—including the prospect of dismissal—as a substantial reason behind the improvements.

Another study by researchers Eric Taylor and John Tyler in Cincinnati Public Schools found that mid-career teachers were more effective at raising student achievement during the year they were evaluated than in prior years.[18] And they were even more effective in the years to follow. Although the researchers could not isolate the cause of the improvements, the evaluation system included four observations by trained principal and peer observers, which culminated in feedback to the teacher observed, and tied results to tenure decisions.

In addition, surveys of Tennessee's teachers show that it is possible for educators to become more comfortable with evaluation systems and see them as mechanisms

for improving their practice and student learning.[19] From the 2012–13 to 2014–15 school year, the proportion of Tennessee teachers who thought the evaluation system improved teaching rose from 38 to 68 percent.[20] Likewise, the proportion of teachers who agreed that the system improved student learning increased from 28 to 63 percent.

> While new teacher evaluation systems are still relatively young, some evidence exists that investing in better systems for both accountability *and* development could help teachers (and their students) improve.

Thus, evaluation systems that include frequent, accurate feedback by trained observers and signal that poor performance has consequences, can help to improve teacher practice and student learning, particularly over time as teachers become more comfortable with and confident in them. However, while such mechanisms are necessary, they are insufficient for promoting all teachers' development at scale. For example, despite Tennessee's positive trend in teachers' perceptions of evaluations, "fewer than four out of ten teachers [surveyed] say they are provided with adequate time for collaboration or access to instructional resources and expertise" based on their evaluation results.[21] For evaluations to fulfill their potential for bolstering teacher practice, ensuring a strong connection between evaluation data—particularly classroom observation feedback—and targeted professional development (PD) is critical. (See "Can Professional Development Improve Teacher Practice?" on page 9.)

Although schools and LEAs have historically had primary charge of teachers' professional development, states can play an important role in ensuring this connection occurs. The remainder of this paper delves into this role more deeply.

Can Professional Development Improve Teacher Practice?

While there is general agreement around teachers' importance and the need to develop strong teachers for every student, there is some debate over whether—and how—teachers can and do improve their practice. U.S. schools spend millions of dollars on "professional development" (PD) for teachers annually, but many studies of specific PD activities find they have little to no impact on teacher effectiveness or student outcomes.[i] Why? One major factor is that, although teachers have unique strengths and weaknesses, they most often receive "one-size-fits-all" PD, such as one-time workshops, that are less likely to improve the quality of classroom teaching and learning.[ii]

Despite the limitations of much current, formalized PD, research suggests teachers improve significantly during their initial years teaching and are able to continue improving throughout their careers.[iii] Teachers tend to improve the most in supportive school contexts with strong school leadership, opportunities for "productive and sustained" peer collaboration,

and, importantly, fair systems for evaluations and meaningful feedback in place.[iv]

In fact, cognitive scientists say that high-quality feedback is the key ingredient for new employees to gain proficiency and others to gain mastery or expertise across a wide range of professions.[v] Not only is there evidence that teachers can and do improve across experience levels, but teachers desire to improve, too.[vi] The vast majority of teachers crave quality feedback on their practice from individuals they can trust.[vii] Accordingly, ongoing, high-quality, actionable feedback and learning opportunities are important for all, not just new or struggling, teachers, and may have the power to drive improvement at scale where other professional development opportunities have fallen short. But in order to assess the impact of these development efforts, having reliable teacher evaluation data and measures of impact on student learning is critical.

i Tom Loveless, *What Do We know About Professional Development?* (Washington, DC: Brookings Institution, February 19, 2014).

ii *Teachers Know Best: Teachers' Views on Professional Development* (Seattle, WA: Bill & Melinda Gates Foundation, 2015); Jenny DeMonte, *High Quality Professional Development for Teachers: Supporting Teacher Training to Improve Student Learning* (Washington, DC: Center for American Progress, July 2013).

iii Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor, "How and Why Do Teacher Credentials Matter for Student Achievement?" Working Paper 12828 (National Bureau of Economic Research, January 2007); John P. Papay and Matthew A. Kraft, "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-Term Career Improvement," *Journal of Public Economics* 130 (October 2015): 105–119; Helen F. Ladd and Lucy C. Sorenson, *Returns to Teacher Experience: Student Achievement and Motivation in Middle School* (Washington, DC: CALDER Center, March 2014); TNTP, *The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development* (New York: TNTP, 2015).

iv Matthew A. Kraft and John P. Papay, "Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience," *Educational Evaluation and Policy Analysis* 36, no. 4 (January 30, 2014): 476–500.

v Craig D. Jerald, *Movin' It and Improvin' It!: Using Both Education Strategies to Increase Teaching Effectiveness* (Washington, DC: Center for American Progress, January 2012).

vi Suzanne McCabe, "Teachers Want Evaluations That Will Help Them Improve Their Practice," Scholastic.

vii American Institutes for Research, "'Generation Y' Teachers Want More and Better Feedback from Principals and Their Peers, A New Study Finds," press release, April 8, 2011; *Primary Sources: America's Teachers on Teaching in an Era of Change*, Third Edition (Scholastic and Bill & Melinda Gates Foundation, 2014).

STATE POLICIES TO SUPPORT TEACHER EVALUATION AS A DEVELOPMENT TOOL

According to the National Council for Teacher Quality's *2015 State of the States* report on educator evaluation, 31 states have a policy in place, either through statute, regulation, or other guidance, requiring that teachers' professional development be based on their evaluation results.[22] However, these states have made different policy choices along the way that affect their systems' ability to support teacher growth. (See Figure 2 on pages 12 and 13 that outlines the policies those 31 states have in place to support a connection between evaluation and development.) While policies alone cannot ensure a connection between evaluation and professional learning, they can serve as a critical foundation. We have organized these policies into three categories: 1) Frequent Feedback, 2) Accurate Feedback, and 3) Formal Development Structures.

Frequent Feedback

To drive improvement, evidence suggests that teachers need to receive frequent, high-quality feedback on their practice.[23] But, to minimize the burden on observers, 22 of the 31 states do not require **multiple, annual observations** of all teachers. While a reasonable goal, without annual observations, it is likely that many teachers will not receive any targeted feedback during the year. Additionally, in nine of these states, **annual summative evaluations** are not required, and teachers are on up to a three-year evaluation cycle, which could limit the frequency of

feedback that teachers receive unless teachers engage in regular coaching or formative feedback apart from their evaluations. Other components of annual evaluation systems, such as student surveys, can provide teachers with timely development information as well, if conducted multiple times during the school year.

Accurate Feedback

Multiple observations are important not only for timeliness of feedback, but also for accuracy, a key element for ensuring that the feedback provided will be useful, and for promoting teachers' trust in the use of observation data for improvement purposes. In some cases, fewer observations may not negatively impact data accuracy—for example, allowing consistently highly-rated teachers to opt out of some observations, as does Louisiana—and can free up observers' time to support those teachers who need it most. However, in general, research shows that having at least two observations per year increases the reliability of evaluation data.[24] Additionally, with more observations, teachers are more likely to feel that their observer(s) saw a fuller picture of their practice.

But for observations to be truly accurate *and meaningful*, those tasked with observing teacher practice must receive *ongoing* **high-quality training and certification**. Almost all—30 of 31—have a

requirement that observers receive training on how to accurately and reliably rate the various components of teacher practice, while only 13 require that observers also become "certified" by passing an assessment toward that end, typically following the training. What nearly all states have in common is that their primary focus for both training and certification has been on rating accuracy. An exclusive focus on ratings falls short of ensuring that observers are also capable of providing constructive feedback and aligning that feedback with targeted development opportunities. Raising the bar for observer certification and requiring that training and certification be ongoing will help ensure observers are equipped to do both.

Only four states require that some or all teachers be evaluated by **multiple observers**, which research shows can increase the accuracy and reliability of observation ratings if the observers are well-trained. Having more than one observer can also boost the content- and grade-level-specificity of feedback, if at least one is an expert in those areas.[25] And this policy can help reduce the burden placed on school principals who often lack the time to conduct multiple observations of all teachers. While about half of the states *allow* for multiple observers to inform teachers' summative assessments, few LEAs actually take advantage of this flexibility, so one observer is often common practice even in those states.[26]

Just as aggregating feedback from multiple trained observers can enhance an evaluation's accuracy, reliability, and formative value, so too can aggregating feedback across multiple students through the use of **student surveys**.[27] Students spend more time "observing" and interacting with teachers and their instruction than any outside observer possibly could. As a result, they can provide valuable insights. Student feedback from developmentally-appropriate

and carefully designed surveys, such as Tripod, have been shown to correlate with and even predict student achievement gains.[28] Over half of the 31 states allow LEAs to use student surveys as part of teacher evaluation, but only five require them.

Formal Development Structures

Observers also need to analyze and leverage evaluation data in order to guide teachers' professional learning. To facilitate this connection, it can be useful to have structures embedded in the evaluation process. **Post-observation conferences** are one potential structure to carve out often hard-to-find time to discuss observation feedback and areas for growth while it is still fresh. **Growth plans**, where observers and/or educators identify areas for development, set professional learning goals, and suggest or select aligned professional learning opportunities, are another. However, while about half of the 31 states require a growth plan for all teachers, the other half only require them for struggling teachers, and often label them "improvement plans." This exacerbates the perception among teachers that "improvement" is something that one is singled out for only when not performing well.[29]

While policy design can lay the groundwork for high-quality feedback and connected professional learning to occur at the local level, policy alone is insufficient. For example, requiring that teachers have a professional growth plan informed by their evaluation results is meaningless unless LEAs ensure these plans are relevant to the most pressing needs of individual teachers and their schools and are acted upon. Although it is ultimately up to LEAs to carry out these policies, states play several vital roles in ensuring they do so, which we discuss in detail in the following section.

> While policies alone cannot ensure a connection between evaluation and professional learning, they serve as a critical foundation.

Figure 2: Policies to Support a Connection between Evaluation and Development Among 31 States with Policies Requiring Evaluation to Inform PD

	Frequent Feedback		Accurate Feedback				Formal Development Structures	
	Annual Evaluation Required	Multiple Annual Observations Required for ALL Teachers	Observer Training Required	Observer Certification Required	Multiple Observers Required for Evaluation	Student Surveys Required	Post-Observation Conference Required	Growth Plan Required
AZ	Y	N	Y	N	NS	A	N	S
AR	Y	N	Y	Y	A	A	Y	E
CO	Y	N	Y	N	A	A	N	E
CT	Y	Y	Y	Y	NS	A[1]	Y[2]	S
DE	Y	N	Y	Y	A	NS	Y	S
DC[3]	Y	Y	Y	N	R	A	Y	NS
FL	Y	N	Y	N	A	A	N	S
GA	Y	Y	Y	Y	NS	R	N	S[4]
HI	N	N	Y	Y	NS	R	Y	E
KY	N	N	Y	Y	A	R	Y	E
LA	Y	N[5]	Y	Y	A	NS	Y	E
ME	N	N	Y	N	A	NS	Y	E
MA	Y[6]	Y	Y	N	A[7]	R	N	E
MI	Y	N	Y	N	A[8]	NS	N	S
MN	N	N	Y	N	A	A	N	E
MS	Y	Y	Y	N	A	A	Y	S
NJ	Y	Y	Y	Y	A[9]	NS	Y	E
NM	Y	Y	Y	Y	A	A	Y	S
NY	Y	Y	Y	Y	R	A	N	S
NC	Y	N	Y	N	R[10]	A	Y	E
ND	Y	N	Y	N	NS	A	N	NS
OH	Y	N	Y	Y	A	A	Y	E
OR	N	N	N	N	NS	A	Y	E
RI	N	N	Y	N	A	A	N[11]	S
SC	N	N	Y	N	R	A	Y	E

	Frequent Feedback		Accurate Feedback				Formal Development Structures	
	Annual Evaluation Required	Multiple Annual Observations Required for ALL Teachers	Evaluator Training Required	Evaluator Certification Required	Multiple Observers Required for Evaluation	Student Surveys Required	Post-Observation Conference Required	Growth Plan Required
SD	N	N	Y	N	NS	NS	N	S
TN	Y	Y	Y	Y	NS	A	Y	E
UT	Y	N	Y	Y	NS	R	Y	S
VA	N	N	Y	N	NS	A	N	S
WV	Y	N	Y	N	NS	NS	Y	S
WY	Y	N	Y	N	NS	A	N	NS
Total	22–Y	9–Y	30–Y	13–Y	15–A, 4–R	19–A, 5–R	18–Y	14–S, 14–E

Legend:
Growth plan indicators: S–some teachers (struggling or novice), E–every teacher, NS–not specified
Multiple observers indicators: A–allowed, R–required, NS–not specified
Student surveys indicators: A–allowed, R–required, NS–not specified
All other indicators: Y–yes, N–no

Note: In this chart, the data under the categories "Frequent Feedback" and "Accurate Feedback" are based on NCTQ's *2015 State of the States* report, the *2015 State Teacher Policy Yearbook*, and consultation with the state education agencies. The data under the final category "Formal Development Structures" are based on New America's review of state legislation and regulation as well as consultation with the state education agencies.

[1] Connecticut bases 5% of evaluation scores on whole-school student learning indicators OR student feedback.

[2] Connecticut requires mid- and end- of year conferences with all teachers, during which evaluations are discussed.

[3] NCTQ's paper and New America's research examine the policies of the Office of State Superintendent of Education (OSSE), the state education agency for the District of Columbia, not DC Public Schools (DCPS).

[4] Georgia is piloting growth plans and will require them for all teachers in the 2017–18 school year.

[5] In Louisiana, all teachers received multiple observations from the 2013–14 through the 2015–16 school year, during which time the state paused the value-added measure in evaluation—the basis for highly rated teachers waiving some observations.

[6] Massachusetts LEAs must annually submit evaluation ratings to the state for every licensed educator. While some years are marked as "formative," ratings are still collected by the state.

[7] Massachusetts' Peer Assistance and Review (PAR) program, specified as an allowable LEA framework for evaluation, employs multiple observers for evaluation.

[8] Other trained observers are allowed beginning in the 2016-17 school year.

[9] New Jersey requires multiple observers for all non-tenured teachers and all teachers on a corrective action plan.

[10] North Carolina requires multiple observers for probationary teachers only.

[11] Rhode Island requires beginning-, mid-, and end-of-year conferences between teachers and their observers.

STATE ROLES BEYOND POLICY: COMMUNICATION, SUPPORT, AND MONITORING

Three state-level strategies can help ensure evaluation systems inform all teachers' professional learning, at the individual and collective levels: 1) clearly **communicating** that evaluation, particularly the classroom-observation component, is a tool for educator development and providing messaging tips for LEAs to employ and reinforce to schools toward that end; 2) offering **support** to LEAs and schools for making meaningful connections between evaluation and development; and 3) **monitoring** the implementation of evaluation systems to ensure they provide accurate, actionable feedback aimed at informing teacher growth. While each strategy is an important step toward meeting this objective, all three are requisite for success.

In our interviews with states that have a policy requiring evaluation to inform teachers' professional growth, individuals at the state education agencies (SEAs) explained their communication, support, and monitoring strategies. Below is a representative summary and analysis of what they shared with us.

Communication

In order for teachers to engage with evaluation systems in ways that will lead to positive changes in their practice, they must not only believe in the systems' fairness, but that their primary purpose and value is to help students meet their full potential by supporting teachers' ongoing development. Ensuring that the purpose is clearly and consistently messaged, directly and indirectly, starts at the top, from the state level.

Most states communicate the need to use teacher evaluation to support teacher growth and learning through online guidance and policy documents, SEA websites, and newsletters to LEAs and educators. For instance, the Massachusetts Department of Elementary and Secondary Education (ESE) messages its evaluation process, which includes professional growth planning and formative observations for all teachers, in a suite of online "quick reference guides" on its website and in educator newsletters.[30] ESE has also used its newsletter to highlight a new video series that includes interviews with administrators, school leaders, and teachers across four LEAs on how the evaluation system can be used to drive educator growth at the local level.

Of course, simply inserting professional-growth-oriented language in guidance documents and newsletters is insufficient for ensuring that educators internalize the message that evaluation is a tool to support teacher development. And such efforts are low-hanging fruit. That is, any state with a teacher evaluation policy could and should implement these communication strategies almost immediately. However, as with many areas of education, the easiest work to undertake is also the least likely to make an impact. Other efforts, while requiring more time, capacity, and investment to implement, are likely to elicit a bigger impact on teacher feedback and, in turn, teacher practice.

Some states strive to create two-way communication channels. For instance, Massachusetts created educator advisory cabinets to provide input on various ESE initiatives, including educator evaluation and support.[31] And during initial implementation of its evaluation system, Kentucky partnered with the Southern Regional Education Board and Hope Street Group to obtain teacher feedback through focus groups and surveys. The state also sought input from its evaluation system steering committee, comprised of representatives from the state association of school administrators, school board association, and teachers union. After finding that some teachers had misconceptions about the purpose of the evaluation system, the state adjusted its communication strategy.[32] Rather than continuing to send newsletters solely to LEA and school leadership, Kentucky initiated a separate newsletter for teachers, and ensured both newsletters' content discussed how evaluation was related to other key initiatives, such as teacher leadership and educator equity.

Beyond engaging in two-way communication with educators, Kentucky messages to LEAs and school leaders more indirectly by providing access to its Continuous Improvement Instructional Technology System (CIITS). The system houses evaluation data along with standards-aligned instructional materials, formative assessments, and professional learning resources and opportunities.

Christine Meisberger, Manager of the Kentucky Department of Education's Teacher and Leader Effectiveness Branch, said that the state strives to use CIITS' comprehensiveness to communicate that evaluation, curriculum, and professional learning must be approached as one coherent entity in the service of teacher practice and student learning, although the extent to which these efforts have paid off remains unclear.[33] Kentucky is in the process of coordinating among its evaluation, professional learning, and curricular and assessment teams to create clear, consistent messages regarding evaluation's key role in supporting educator growth to teach to the Kentucky Academic Standards.[34] While this type of coordination is often difficult for SEAs, it will likely come easier to KDE since, unlike in most other SEAs, all of these teams are housed under one banner: the Office of Next-Generation Learners.[35]

Kentucky's CIITS is not only an indirect messaging strategy but also an online support tool for LEAs to use at their discretion. Other states have adopted similar electronic platforms to support the connection between evaluation and professional learning at the local level, as highlighted below.

Support

Unless states go beyond communication to support LEAs and schools in making the link between evaluation and professional learning, many at the local level will have difficulty following through. State education agencies can support LEAs and schools by providing them with clear guidance, tools, as well as hands-on support and capacity-building in connecting evaluation results to development opportunities.

Most states provide resources to help teachers and their observers understand the classroom observation rubric and other evaluation measures, a prerequisite for using evaluation to appropriately guide professional growth. Most also attempt to provide some specific tools and implementation guidance around these measures. For instance, since Massachusetts allows for student surveys as an additional measure in evaluation and source of formative feedback, the state provides LEAs with validated model instruments and administration protocols to aid implementation.[36] To guide the educator growth plan process, Massachusetts also equips LEAs with a protocol for developing professional learning goals as well as tips for how best to use evaluation data to inform this learning at the LEA, school, and individual teacher levels.[37]

Some states have gone beyond providing guidance and have invested in face-to-face support to help LEA leaders make the link between evaluation and development. Arizona, for instance, recently held a two-day summit on using evaluation to support professional learning.[38] The sixth of its kind, the summit brought together over 300 LEA leaders for sessions including "Evaluation Feedback for Professional Growth," "Principals Coaching Teachers," and "Creating a Culture of Data Use and Professional Learning." When states invest in face-to-face support strategies like these, they must ensure they are high-quality and reach as many educators as possible. Additionally, while the research on professional learning finds that one-time seminars or workshops

are not very effective for ensuring knowledge transfer and follow-through, those that involve participants in active and application-based learning can prove meaningful.[39] States should make concerted efforts to provide LEAs with ongoing supports grounded in best practices for adult learning and encourage LEAs to do the same for school leaders.

In addition, many states use online support tools to promote a connection between evaluation and professional learning. Several, including Delaware, Connecticut, and Arkansas, have contracted with BloomBoard to provide an online platform with PD resources that correspond to specific teaching practices in the states' observation rubrics. Observers can recommend or educators can select resources based on observation results. About half of LEAs in Delaware and Connecticut, and all in Arkansas, use the BloomBoard system.[40]

> **Online tools and resources can be valuable in providing educators a clear path for professional growth on specific teaching practices and content, but may be underused if educators are not aware of or confident about how to use them.**

Several other states, such as Kentucky with CIITS, have also developed their own online or "blended"—online and in-person—professional learning modules that are aligned to their evaluation systems' observation rubrics. States use different approaches for developing and vetting these resources. Kentucky, for instance, uses a tiered process where teacher-created resources are vetted first by the school, then by the LEA, and finally by the state before being uploaded into the system for widespread use. And Georgia created an online system of SEA-developed professional learning modules that cover formative instructional and assessment practices. The modules help build educators' skills in implementing the state's new academic standards and cover practices included in the observation rubric, such as collecting and analyzing evidence of student learning.[41]

After examining trends in teacher evaluation data, New Jersey realized its educators needed support in

implementing specific teaching strategies required by new academic standards, including fostering intellectual engagement through effective questioning and discussion strategies and having evidence-based conversations. To address this need, the state partnered with three LEAs to develop blended PD modules that were targeted to those areas of the observation rubric.[42] In addition to making training materials available on its website, the SEA also trained 150 teacher leaders across 19 LEAs to then train their peers on the strategies in person.

Online tools and resources can be valuable in providing educators a clear path for professional growth on specific teaching practices and content, but may be underused if educators are not aware of or confident about how to use them. For instance, in Delaware, very few educators actually use the resources made available through BloomBoard. "People are still getting comfortable with it—both teachers and evaluators. [There are] a limited number of sophisticated users linking [evaluation data] to professional development," explained Eric Niebrzydowski, Deputy Officer of Teacher and Leader Effectiveness at the Delaware Department of Education (DDOE).[43]

Thus, states should go beyond providing access to linked evaluation and development platforms and directly support LEAs in harnessing their potential. For the 2015–16 school year, Arkansas's educator effectiveness team is partnering with the school improvement and professional development teams to support four high-poverty, high-minority LEAs in linking evaluation data with professional growth opportunities via BloomBoard.[44] For up to two years, the state will offer two half-day in-person trainings and on-demand coaching, during which school and teacher leaders will review teacher evaluation data from the prior school year and plan professional learning goals for the current year—both school-wide and for individual teachers.

Even when used well, however, online platforms and modules will be insufficient for supporting the connection between evaluation and development at the local level if those responsible for evaluation are not able to observe teacher practice accurately and provide meaningful feedback. Additionally, while they fill a current void, these tools can only provide a limited scope of professional learning and training materials, which may not meet the needs of all teachers. For instance, online professional learning

modules covering basic teaching strategies will likely only be useful for new teachers or those struggling in particular areas. As a result, observers must be skilled in suggesting or planning aligned learning opportunities beyond those readily available online.

To date, the states that have provided observer training and/or certification have focused more on ensuring observation rater accuracy and less on ensuring observers can provide teachers with constructive feedback and meaningful supports aligned to those ratings to help them improve. But a few states have moved ahead here, including Tennessee and Rhode Island. For instance, in its observer trainings, Rhode Island strives to develop principals' ability to analyze evaluation data to plan aligned school- and teacher-level professional learning.[45] The SEA leads principals in digging into mock evaluation data, brainstorming PD options, and then planning targeted PD based on the results.[46]

Likewise, North Carolina has used federal Race to the Top (RTT) funding to offer biannual "Principal READY" sessions to build principals' ability to analyze evaluation data, provide actionable feedback to

teachers, and coach them toward growth.[47] Because RTT's funding ended last year, the state worked to find other funding to sustain these sessions during the 2015–16 school year because principals found them valuable.

A few states that have chosen to give more control to LEAs in evaluation design have developed networking sessions for educators to share tools and best practices around evaluation and development. (See: "Strict vs. Flexible Evaluation Models" below.) For instance, in the 2015–16 school year, Minnesota is offering monthly networking sessions for peer and school-leader observers to share ideas on how to have growth-oriented conversations and make the most of post-observation conferences.[48] In addition, the SEA offers several scheduled and by-request sessions for school leadership on using data to plan for school-wide, team, and individual professional development. Recently, demand for these has been increasing beyond the state's current capacity to deliver.

Similarly, in the 2014–15 school year, Massachusetts offered professional learning networks (PLN) for up to eight LEAs to partner around building observer

Strict vs. Flexible Evaluation Models

In designing evaluation systems, states have generally adopted one of three models: 1) a statewide system which all LEAs are required to adopt; 2) a state model which LEAs may choose to opt out of if they produce a comparable system; and 3) LEA-designed systems which must adhere to some specific criteria provided by the state (e.g., which measures must be included).[i]

Having flexibility with evaluation design may promote greater investment in implementation at the local level, as educators may have had more opportunity to weigh in and thus feel more ownership of the system. However, most LEAs do not have the technical expertise and capacity to design their own valid and reliable systems. What is more, such flexibility can make it difficult for the state to support and monitor the systems' use, including for teacher development.

For example, classroom observations of educators' practice are typically the largest measure in teacher evaluation systems and hold the most promise for providing teachers with actionable feedback. But in some states, such as Florida, each LEA may choose a different observation framework for rating teacher practice (e.g., Danielson, Marzano).[ii] A multitude of different local systems limits states' ability to provide tools and resources demonstrating what strong practice looks like on each aspect of the observation rubric. It also complicates states' efforts to provide evaluator training to promote accurate ratings and actionable feedback—both key elements in ensuring teachers trust the evaluation process and the feedback they receive within it, and can apply that feedback to their practice.

i Kathryn M. Doherty and Sandi Jacobs, *2015 State of the States: Evaluating Teaching, Leading, and Learning* (Washington, DC: National Council on Teacher Quality, November 2015).

ii Tysza Gandha and Andy Baxter, *Toward Trustworthy and Transformative Classroom Observations* (Atlanta, GA: Southern Regional Education Board, February 2015).

capacity. And beginning in the 2015–16 school year, Kentucky deployed field staff in each region across the state to develop principals' feedback and coaching skills.[49] States should continue to communicate and expand opportunities like these to help ensure that observers know how to provide high-quality feedback to teachers and guide their ongoing learning.

However, solely making such supports available and communicating their existence may not be sufficient to get all LEAs and school leaders to adequately focus on ensuring their feedback is accurate and helps teachers develop. The next section examines how monitoring is an additional tool states can use to this end.

Monitoring 🔍

Several conditions must be present for evaluation to spawn meaningful, aligned professional learning opportunities, and states can monitor their presence using a variety of strategies depending on their context, including level of data access and staff capacity. First, states can ensure observers are well-trained to provide valid, reliable evidence of educators' practice during classroom observations to inform their professional growth. For instance, for its state evaluation model, Tennessee tracks whether observers complete its training and certification program and get recertified annually.[50]

Under ESEA waivers, states were required to have a plan for monitoring implementation of local evaluation systems.[51] (See: "The Rise of New Teacher Evaluation Systems," on page 6.) However, very few states have gone beyond monitoring the basic components of their evaluation systems to ensuring that they meaningfully support LEAs, schools, and teachers in their continuous growth. For example, with the exception of Delaware, no state mentioned auditing the quality of feedback that teachers receive—states generally leave such monitoring up to LEAs likely due to the challenge of doing so from the state level.

States should not need federal oversight to commit to monitor whether LEAs are using evaluation data as tools to bolster improvement in teachers' practice, nor should they need a federal push to target support to LEAs that are not fulfilling this charge. They could fulfill these roles in a variety of ways that make sense for their context. For instance, high classroom observation scores paired with low student growth scores, and vice versa, indicate that there may be a need to support observers in better interpreting

evidence of educator practice, since generally strong teaching should result in improved student outcomes. New York is one of several states with largely locally-determined systems that monitors trends in LEAs' evaluation results, and the correlation between overall observation ratings and measures of student learning in particular.[52] The state then works with those LEAs whose evaluation results are most misaligned to develop a plan for ensuring observation rating accuracy, whether it is suggesting alternative observer training or requiring external independent observers.

States using online platforms linking evaluation and professional learning can also monitor teachers' completion of professional development opportunities made available through the platform. For instance, Georgia analyzes teacher participation in its online modules by teachers' summative evaluation performance level.[53] Arkansas plans to take it a step further with its four-LEA pilot by attempting to track via BloomBoard whether teachers' participation in prescribed resources has an impact on their practice as defined by classroom observation scores.[54] The proposition is, "if teachers are provided with appropriate feedback and high-quality support, and if they authentically engage in the support provided, did they improve performance?" said Ivy Pfeffer, ADE's Assistant Commissioner of Educator Effectiveness.

Finally, states can monitor educators' perceptions of whether evaluation informs their professional growth and the quality of aligned learning opportunities they receive. Several states have undertaken surveys or focus groups on these topics to inform their evaluation and professional learning efforts. For example, with its online modules, Georgia includes a survey at the end of each to help assess users' perception of quality and the overall experience.[55] Beyond state-developed surveys, 20 states administer the annual Teaching, Empowering, Leading, and Learning (TELL) survey and at least seven administer the Standards Assessment Inventory (SAI2) survey. The former, developed by the New Teacher Center, asks educators whether PD is targeted at their needs and aligned with school goals. The latter, developed by Learning Forward, helps states that have adopted the organization's standards for professional learning assess the quality of PD offered in LEAs and whether it actually aligns with the standards.[56] Although these perceptions may be somewhat subjective based on teachers' individual reference points, they can yield additional insights into areas for system improvement.

While a state's ability to monitor will be more or less difficult depending on its size and authority over evaluation, it is imperative that it pay close attention to local implementation in order to know how best to target communication and supports. Still, monitoring strategies like those included above will only have an impact on improving educator practice if states act on the information obtained. For instance, tracking

participation in professional growth opportunities is meaningless unless states also have a plan to track the quality and impact of those activities over time, and intervene where they are low. Likewise, states can engage educators through surveys and focus groups, but the results will be meaningless unless they address issues and concerns.

STATE SNAPSHOTS: FROM ACCOUNTABILITY TOWARD SUPPORT

💬 **Communication**   👥 **Support**   🔍 **Monitoring**

The SEA efforts detailed above, while not exhaustive, represent the types of efforts detailed in our interviews with states that have policies requiring evaluation to inform teachers' professional growth. However, most states acknowledged that to date their initial evaluation system efforts have focused on operationalizing them for accountability rather than on ensuring that the resulting data meaningfully support changes in teacher practice. Few states have yet to take meaningful strides to ensure evaluation is connected to ongoing professional development at the local level. And none have done so in a comprehensive way: from **communication**, to **support**, to **monitoring**.

This should not be surprising. Communicating, supporting, and monitoring the basic components of evaluation systems are a heavy lift for many states in and of themselves. But most states indicated they

want to do more to leverage evaluation as a tool for development and are planning to do so moving forward.

A handful of states offer promising practices and lessons learned that can help inform this future work. Following is a more in-depth look at four states' efforts to re-envision evaluation as a tool for support that could ultimately have staying power among educators. All four—Louisiana, Colorado, Tennessee, and Delaware—were winners of the federal Race to the Top competition and earlier adopters of multi-measure evaluation systems. While some are further along in these efforts than others, all offer insights into how other states, even without competitive funds, can begin making their evaluation systems as much as about development as accountability.

Louisiana: Repurposing Evaluation as a Tool for Accountability *and* Support

In the 2012–13 school year, the Louisiana Department of Education (LDOE) rolled out an ambitious teacher evaluation system: "Compass," short for "Clear, Overall Measure of Performance to Analyze and Support Success." The department intended for Compass to support student success through both teacher accountability *and* support. However, as in other states, the move sparked pushback among many educators who focused primarily on the accountability aspect.[57] Educators saw one aspect of the system as particularly unfair: if rated "Ineffective" on either of the two system components (professional practice or student learning growth), teachers were rated "Ineffective" overall regardless of the score for the other component, which is known as the "Ineffective Override" policy.[58]

The following year, 2013–14, Louisiana also rolled out new, more rigorous standards for student learning and related assessments. In order to be fair to teachers during the transition, LDOE decided to suspend the requirement of using a student growth measure based on state assessments to calculate a teacher's summative evaluation rating through the 2015–16 school year. By doing so, LDOE communicated to educators in tested subjects and grades that it wanted to provide them with more time to grow their practice to teach to the more rigorous academic standards.

However, LDOE continued to produce a state-assessment-based student growth measure for LEAs and evaluators to access for informational purposes and use for stakes at their discretion. Additionally, the automatic "Ineffective override" policy remained in place, which served to continue perceptions that Compass was primarily about accountability rather than supporting educators' growth. Meanwhile, LDOE's communications on Compass and the new academic standards came from separate LDOE teams, reinforcing this perception.

Recognizing these issues, LDOE took further steps to make Compass more of an instructional improvement tool for educators. First, in the 2014–15 school year, LDOE decided to bring its Compass team under the Academic Content umbrella in an effort to ensure educators perceive evaluation as a means for receiving feedback on how to teach to the new more rigorous

standards.[59] Though this change happened recently, in the long run such a structural change is likely to have an impact on educator perceptions' of evaluation as supporting their daily teaching practice.

All LDOE communications materials and resources related to Compass now make a connection to the standards. "In all that we do, we're trying to make 'Compass' more about providing teachers feedback and support on teaching the new academic standards," said Rebecca Kockler, Assistant Superintendent for Academic Content.[60] Such resources include new math and English Language Arts observation and feedback guides with key "look fors" in each subject and a principal guidebook for teaching and learning.[61] The latter walks school leaders through the process of choosing curricula and assessments, monitoring educators' progress on observations and feedback, and planning for targeted professional learning within a system of school-wide instructional support.

> **The department encourages principals to tap all teacher leaders to provide ongoing low-stakes, content-specific feedback that helps teachers transition to the new academic standards.**

In addition, LDOE brought a committee of educators and community stakeholders together to recommend changes to Compass, which the State Board of Education approved. Based on these recommendations, LDOE removed the "Ineffective Override" policy beginning in the 2015–16 school year and allowed school leaders more autonomy in making personnel decisions. Principals may now use their discretion to adjust the test-based student growth measure by one level in either direction. In making this change, the department sought to communicate that evaluation systems "must empower principals and give them the tools to improve teaching and learning."[62]

Since the summer of 2013, LDOE has also hosted an annual Teacher Leader Summit for both teacher leaders and principals to reinforce the message that the observation process is as much about growth as it is about accountability, and to provide them with the skills to make good on that message. Although principals are primary observers, LDOE allows them to designate individuals in formalized roles (e.g., master teachers, coaches) to be trained and serve as additional observers to inform their final ratings and feedback. In addition, the department encourages principals to tap all teacher leaders to provide ongoing low-stakes, content-specific feedback that helps teachers transition to the new academic standards.

During the 2015 summit, which New America observed, principals and teacher leaders attended joint sessions on how to deliver actionable feedback and coach teachers towards growth. Together, they watched videos of teachers executing lessons and collected evidence using expanded content observation and feedback tools with key "look fors" in ELA and math. They then brainstormed next steps for the teacher to take along with appropriate, aligned supports.

Still, getting teachers to view evaluation as a tool for support will require further effort by the state. At the 2015 Teacher Leader Summit, a small group of teacher leader advisors New America interviewed said they appreciated LDOE's efforts to make evaluation more about development and generally felt support in implementing the new academic standards.[63] However, in order to better focus their improvement efforts, the group wanted to receive more frequent and higher quality feedback on their practice, and cited school leader development and observation frequency as critical needs for the state to address.[64] In addition, the group valued the opportunity to set professional growth goals but found the required development plans to be more compliance-driven than meaningful, likely due to to a need for stronger school leader training and guidance.

Summit opening session in 2015. Photo: Courtesy of the Louisiana Department of Education.

In order to deliver on these areas, and its stated goal of providing principals tools to improve teaching and learning, the state should focus on building principals' and LEAs' capacity to support teachers in two areas: 1) connecting teachers to development opportunities and resources aligned to areas for improvement; and 2) providing frequent high-quality feedback, by leveraging teacher leaders and other means.

LDOE has already begun some positive changes in these areas. For the 2015–16 school year, the SEA will provide a yearlong fellowship for school leaders to build their skills in instructional leadership, including how to help teachers set goals, provide them with meaningful feedback, and facilitate supports based on their evaluation results.[65] While the fellowship will reach only 10 percent of Louisiana's school leaders in its first year, LDOE hopes to scale up the work over time. And the state will expand its funding to LEAs for the alternative TAP evaluation model, which provides teachers with more frequent observations and aligned supports, including weekly team meetings with mentor teachers by grade or subject.[66]

Colorado: Moving Beyond Communicating Evaluation as a Tool for Professional Growth

The Colorado Department of Education (CDE) made a point to communicate the importance of using evaluation to support teacher growth during the initial stages of system implementation. Twenty-seven LEAs participated in a pilot of the evaluation system in the 2012–13 school year and implemented it fully through a staged-in process in the following years. In these early stages, CDE partnered with various groups, including the teacher and principal unions and school board association, to develop a list of key messages for LEAs to use for their educator effectiveness efforts. Through these messages, CDE communicated that evaluation was "to provide all educators with more meaningful feedback and support so they can achieve maximum results with students."[67]

There is some evidence that CDE's messaging paid off. In 2014–15, the first year in which the majority of LEAs had implemented their new evaluation system, a TELL survey found that over half of teachers felt their system helped them improve.[68] However, other survey data showed that teachers' confidence in the system's fairness was correlated with how they rated their school leaders, no doubt related to the quality of their observations and feedback.[69]

As such, CDE is now moving beyond messaging to offer online support tools that can help LEA and school leaders provide accurate observation ratings and connect them with professional learning. CDE offers Elevate, an online inter-rater reliability tool, along with RANDA, an online performance management system that stores educators' evaluation data.[70] Both are available for free to all LEAs using the state model system, about 96 percent of LEAs in Colorado. Through Elevate, observers can choose to watch and score videos of instructional practice to improve their rating consistency, although they are not required to gain formal observer certification as in some other states. And RANDA, used by about two-thirds of Colorado LEAs, has the capacity to link to online professional learning platforms that LEAs are already using, such as PD 360.[71] The system also prompts observers to make optional recommendations for professional development when they enter evaluation data.

Not all of CDE's supports to LEAs and schools in this area are online: like many states, it often partners with regional service centers to deliver observer trainings. (See: "The Role of Regional Education Service Centers in Supporting Teacher Development" below.)

Still, CDE recognizes that it has a long way to go in developing RANDA's PD component, including building its capacity to link to other online PD providers and resources and then mapping those resources to teaching practices in the observation rubric.[72] Because the state does not have access to evaluation data at the teacher practice level, CDE has emphasized the value of the platform's data reports that highlight teachers' strengths and weaknesses at the individual, school, and LEA levels. "Rather than a single score, teachers can see an analysis of what brought them up and down [across all teaching

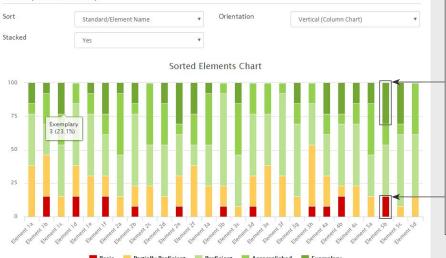## The Role of Regional Education Service Centers in Supporting Teacher Development

Many states rely on their regional education service centers to deliver observer training and other professional development, though the degree to which they rely on them varies by the state's geography and funding authority. For instance, because Colorado contains mostly small, rural, widely dispersed LEAs, the state adopts a "train the trainer" model and partners with approved training providers, including regional centers, to deliver ongoing observer training. While the centers operate independently, CDE uses a rubric to approve the centers' trainings and works with them if they fail to meet CDE's standards.

Most states' regional centers operate entirely independently. That is, LEAs pay to participate in the services provided. So while larger states—or those with geographically isolated LEAs—can leverage these centers to help LEAs connect evaluation and PD, some state education agencies have little authority over center goals or monitoring outcomes for impact.[i] Colorado is one example of how an SEA can leverage regional centers in support of teacher development despite lacking legal authority over them and more states can and should think creatively to do the same.

i Danielle Gonzales, email to authors, January 11, 2016.

Figure 3: Sample RANDA Data Report (School-Level)

With these reports, school principals can click on the performance categories within a specific teaching practice (called "elements") to produce a list of the teachers in each of those categories, along with their grades and content areas. From there, they can easily pair higher- and lower-performing teachers on each element for peer-to-peer professional learning, such as collaborative planning and informal observations.

Source: Courtesy of the Colorado Department of Education

practices], which drives their professional growth plan," said Toby King, CDE's Deputy Director for Educator Effectiveness. Additionally, school leaders can view school-wide or grade-level reports to inform their professional development decision-making, such as pairing teachers for peer observation based on their evaluation results (see: Figure 3 above).

While CDE is providing state funds for RANDA and Elevate, it has neither the funding nor the staff capacity to successfully implement all its desired education initiatives at once.[73] The state ranks in the bottom half of states in terms of both federal and state funding for education resources and full-time staff.[74] In order to expand its support for LEAs in connecting evaluation and development, CDE has partnered with the Colorado Education Initiative (CEI), a nonprofit foundation in the state.[75] One example where CEI funding is helping to build CDE's capacity to promote teacher development is via an annual professional learning summit. Now in its third year, the CEI and CDE summit brings together educators from across the state for sessions around the integration of academic standards, evaluation, and professional learning. For instance, during the 2015 summit, which New America observed, school leaders attended sessions to learn effective feedback strategies employed by the Educator Effectiveness Liaison Network.

Funded by CDE, the liaison network was first implemented in the 2014–15 school year in lower-capacity LEAs to build observers' skills in providing effective feedback grounded in the new academic standards. CDE provided funding for 20 percent of LEAs to select liaisons to receive ongoing training on how to deliver meaningful feedback to teachers and coach them towards growth. Within their LEA(s), the liaisons then provide one-on-one principal professional development and coaching around high-quality observation and feedback on a monthly basis. In particular, they conduct co-observations with a subset of school leaders (determined by the LEA), compare their teacher practice ratings for consistency, and then practice delivering effective feedback during post-observation conferences.

Colorado has put some structures in place to support evaluation as a tool for actionable feedback and clearly communicate that mission. Still, there is more work to be done to support LEAs and school leaders in the process. In the future, CDE should work to ensure all school leaders are capable of analyzing evaluation data to recommend professional learning—perhaps by scaling up its liaison network. The state should also ensure school leaders understand the instructional shifts required by the new academic standards and empower LEAs to create the time and structures for teachers to collaborate. By doing so, CDE can meet its goal of "trying to get away from the idea that you need to go purchase external PD" and instead use RANDA's data reports to guide school and teacher professional learning in support of new academic standards-aligned instruction.[76]

Tennessee: Building Trust as a Foundation for Teacher Development Efforts

In 2011–12, Tennessee became the first state to implement a multiple-measure evaluation system based partly on classroom observations and partly on student growth, a year before teachers were to begin incorporating new academic standards for math and English Language Arts, and two years before LEAs were required to use evaluation results to inform personnel decisions. From the outset of instituting its statewide model, TEAM, the Tennessee Department of Education (TDOE) has prioritized building educator trust in the system.

One way TDOE helped build trust in TEAM was by supporting observers' ability to provide teachers with accurate and reliable data through a rigorous observer training and certification program. Unlike some states' observer training, TDOE's trainings go beyond ensuring rating accuracy. They build observers' capacity to provide teachers with accurate and meaningful observation feedback. Observers spend time watching videos of a teacher's instruction, collecting evidence, scoring the videos, and planning for what they would discuss with the teacher. After the training, TDOE requires prospective observers to pass a certification test, which includes accurately scoring observations and drafting teacher professional learning plans based on the results. The plans must include areas for growth along with aligned suggestions for how to develop them. All observers must pass the test before they can evaluate, and returning observers must be recertified annually through either in-person training or a separate re-certification test.

The state then closely monitors the correlation between educators' classroom observation, student growth, and summative evaluation scores. In the 2012–13 school year, the state began identifying the schools most misaligned on the student growth and observation components of the system in each of the eight regions across the state.[77] Using a combination of federal grant and state funds, Tennessee hired eight TEAM coaches to build the capacity of observers in each of the identified schools to rate teacher practice accurately, provide effective feedback to teachers during post-observation conferences, and coach them towards growth. Activities include monthly co-observations of classrooms to discuss evidence, role-playing coaching conversations, delivering effective

feedback, and analyzing school- and teacher-level evaluation trends to drive professional learning.[78]

In another effort to build trust, TDOE has invested heavily in soliciting and incorporating teacher input on the state model evaluation system. For instance, in 2012, TDOE partnered with the nonprofit SCORE to conduct an independent report of educators' perceptions on evaluation.[79] Since 2012, TDOE has also conducted its own annual survey of teachers' perceptions of TEAM in partnership with the Tennessee Consortium on Research, Evaluation, and Development at Vanderbilt University.[80] The initial survey focused on fidelity of TEAM implementation, including questions on how frequently observations occurred, by and for whom, and the timeliness of the observation feedback. In 2013 and 2014, after a couple years of implementation, TDOE added survey questions on the content, quality, and frequency of feedback teachers receive under TEAM, along with whether observers provided useful suggestions for improvement and then followed up with teachers to determine progress.

Nate Schwartz, Chief Research and Strategy Officer, says that Tennessee educators are increasingly seeing evaluation to be more about improvement rather than judgment, likely due to the state's investment in building observer capacity to provide teachers with trustworthy evaluation data.[81] The survey results support this observation: the number of teachers who felt as though evaluation improved their craft increased from 54 percent in the 2013–14 school year to 68 percent in 2014–15.[82] In 2015-16, 81 percent of teachers thought the evaluation system helped them know the precise areas where they could improve.[83] Another potential reason for positive educator perceptions is that teachers witness TDOE being responsive to its teacher survey outcomes, and realize that the SEA values their input. Rising teacher survey response rates every year provide evidence that this may be true.[84]

Still, TDOE realized that while it had been strong on implementation fidelity, it needed to do more to connect evaluation with meaningful development opportunities.[85] Building on the foundation it has in place from its implementation and data quality efforts,

the SEA is developing supports to help teachers further improve.[86] One way Tennessee is doing this is through a teacher-pairing program called the Instructional Partnership Initiative, developed with researchers from Brown University. Unlike in some other states, TDOE has access to practice-level classroom observation data for individual teachers. Given this, TDOE came up with the idea of generating reports to show principals how they could use the data to pair higher- and lower-performing teachers based on areas of strength and growth in particular teacher practices. "There are all kinds of mentoring structures within schools but they are rarely set up by [teachers'] strengths and weaknesses," Schwartz said.

Instructor development in 2015. Photo: Courtesy of the Tennessee Department of Education.

The paired teachers receive guidance from their principal on how they could collaborate to improve their skills in the area identified, such as reviewing formal observation results together and observing each other's practice to provide informal feedback on the particular matched skill.[87] Initial results from the pilot program appear promising.[88] In comparison to a control group that did not receive the intervention, schools that implemented the partnerships saw student test scores increase by five to six percentage points in math and reading.[89] Participating teachers found their evaluation feedback to be more about development than judgment and appreciated the

opportunity to collaborate and reflect on their practice with peers.[90] While piloted in only one LEA in 2013–14, TDOE expanded the program to include about 20 LEAs in 2014–15, and half of schools in the state in 2015–16. TDOE hopes to make the low-cost program statewide soon.[91]

TDOE has also used evaluation data to identify teachers to lead regional professional development and to assess the impact of the trainings. Beginning in 2012, the state leveraged its evaluation data to identify and train Highly Effective teachers to become "Core Coaches" and provide "train the trainer" support for about half of its teaching force in implementing new academic standards-aligned instruction.[92] In a state-led evaluation of the program, participants showed marked improvements on standards-aligned instructional techniques, including effective questioning and problem-solving support, as measured by classroom observation scores.[93] Student outcomes also improved, and the program has been well received by teachers.[94] In fact, the program likely bolstered teachers' positive perceptions of evaluation and trust in the system.

In the 2015–16 school year, Tennessee also revised its school leader evaluation system to acknowledge principals' role in ensuring a connection between teacher evaluation and development.[95] The system now assesses how well principals implement the teacher evaluation system to inform professional learning goals, engage staff in differentiated professional learning, and deploy teacher leaders to coach their peers. The state recognized that when it comes to driving teachers' professional learning, "principals cannot do it alone and need to leverage highly effective teachers," said Paul Fleming, Assistant Commissioner of Teachers and Leaders.[96]

While promising, the school leader evaluation system will take additional time and training to implement well. Though the state plans to scale up its pilot program pairing teachers based on their strengths and weaknesses, it does not yet have a clear timeframe for extending it to all LEAs.[97] And beyond encouraging paired teachers to meet during regularly scheduled hours—for instance, a planning period or before school—the state has not provided further guidance to LEAs around carving out time for this collaboration.

Delaware: Monitoring to Ensure High-Quality Feedback Informs Professional Learning

After Tennessee, Delaware was the second state to implement a multiple-measure evaluation system under RTT. The Delaware Department of Education (DDOE) fully implemented its system, called DPAS-II, in the 2012–13 school year, the same year the state began fully implementing new academic standards. Since that time, the department has closely monitored LEAs' progress in DPAS-II implementation by visiting a subset of schools across all LEAs throughout the year.[98] During the process, which New America observed, department officials sit down with observers to review the quality and completeness of their observation ratings and feedback entered into the state's online system. In this way, the department seeks to ensure school leaders are conducting observations and entering their feedback as they should. Additionally, the monitoring team examines a sample of evaluations to dig deeper and assess feedback quality. For instance, the team measures whether the feedback is accurate and actionable, hones in on specific areas for growth, and offers suggestions for improvement.

> **During some monitoring visits, department officials glean whether observers adequately coach teachers, offer suggestions for improvement, and point them to other professional learning opportunities.**

During some monitoring visits, department officials also ask teachers to describe an evaluation experience from the current year. They use a series of questions to glean whether observers adequately coach teachers, offer suggestions for improvement, and point them to other professional learning opportunities. The state then compiles their findings in a detailed report for each LEA with suggestions for improvement in the areas identified. For instance, the state may recommend that observers attend supplemental trainings, which the state provides throughout the year. In addition to requiring initial training for new

and returning observers, the DDOE hosts topical seminars and a mid-year training for observers based on areas for growth identified in its monitoring visits, such as post-observation conferences.

Like Tennessee, Delaware does not simply monitor LEAs without also providing some support to help them make the connection between evaluation and professional learning at the local level. The state mandates that all LEAs use an online platform to house and share evaluation data and provides BloomBoard for free as one option. Most LEAs have signed on to use the BloomBoard platform, which, as previously noted, connects teacher evaluation data with a menu of aligned professional learning resources. Delaware then monitors teachers' participation in the professional learning opportunities available through the platform. After realizing that initial take-up among educators was low, it decided to host "BloomBoard University" to provide training for school leader observers, teachers, and LEA leaders on how to maximize the platform to help guide professional growth.[99]

One area the state has struggled with is differentiating teacher performance, with most teachers being rated in the top category even within specific teaching practices of the observation rubric. As a result, teachers "have nowhere to grow," making attempts to target professional learning difficult, Christopher Ruszkowski, Chief Officer of the Teacher & Leader Effectiveness Unit said.[100] Ruszkowski attributes this challenge to the potential need for rubric design tweaks and more rigorous observer training.

Though its initial trainings are required, Delaware's ongoing trainings are optional, and only about five percent of observers attend.[101] In the future, DDOE should consider making some of its ongoing trainings mandatory, particularly for new observers, and clearly communicate the need to connect evaluation and professional learning in order to ensure teachers receive high-quality feedback on their practice and suggestions for targeted support. And since DDOE does not always interview teachers about whether their observation feedback is useful and connected to targeted professional learning, the department should consider doing so in order to track improvements over time.

Common Themes and Considerations

These four states are actively working to forge a better balance between accountability and teacher development in service of student learning. They are going beyond basic messaging to focus on areas like providing observers with training for suggesting specific supports connected to evaluation feedback, and developing methods for meaningfully monitoring LEAs to ensure they are following through in making this connection on the ground.

Building strong school leaders is key to ensuring that the evaluation and development enterprise leads to targeted feedback and professional learning for teachers at the individual, group, and school levels. All of these states have acknowledged the importance of instructional leadership for evaluation's success by making efforts to ensure more leaders have the skills to accurately observe teacher practice and provide high-quality feedback tied to their observations.

However, no state has yet put promising efforts in place to communicate, support, and monitor the connection between evaluation and development. This is partly because just implementing the basic structures of new evaluation systems has been challenging. SEAs also face competing demands on their time, such as implementing new school accountability systems and college- and career-ready standards. No wonder many are just now beginning to think about how to use the data resulting from these systems for any purpose—accountability <u>or</u> development.

Still, it is precisely because states are simultaneously implementing several major new initiatives that the focus on development is so important. For example, new student academic standards require new teacher

knowledge and practices, and a recent Gallup poll found that six in ten teachers believe that these new standards can facilitate better teacher development. However, the poll also found that teachers are fearful of being held accountable for student performance on new aligned state assessments.[102] SEAs appear to be responding to this fear by keeping discussions of evaluation largely separate from standards. Although other states may be doing so, only Louisiana spoke intentionally about working to reframe evaluation as a way to support teacher adoption of the standards.

Capacity constraints have also limited states' ability to scale up the initiatives they have undertaken. For example, while several states provide trainings for observers on how to provide meaningful feedback to teachers and suggestions for growth, most states still struggle to ensure that all leaders are equipped with these skills. Financial resources have played a significant role. Although states had to submit a sustainability plan as part of their Race to the Top applications, the program's end has meant that many are struggling to maintain the initiatives they have instituted even at their current scale. States do need sustained resources for some of this work, particularly when it comes to areas such as ensuring that all new observers are adequately trained to deliver accurate feedback tied to meaningful support.

However, a recent analysis of state spending under RTT found that state grantees spent the least of their funds—nine percent, overall—on educator evaluation and support systems.[103] And two of the states highlighted, Louisiana and Colorado, received a small amount of RTT funds in comparison to other states.[104] As such, these snapshots demonstrate what SEAs can accomplish when they choose to make evaluation and support system improvements a top priority.

> **Building strong school leaders is key to ensuring that the evaluation and development enterprise leads to targeted feedback and professional learning.**

ACTIONS STATES CAN TAKE TO BETTER CONNECT EVALUATION AND DEVELOPMENT

States have historically played a minimal role in teachers' professional development, with schools and LEAs leading the bulk of this work. The unsurprising outcome is that many states, even those with policies requiring their teacher evaluation systems to inform development, have focused on their systems' design and implementation for accountability while largely ignoring their potential for helping teachers improve their practice.

Still, several states—including those highlighted in our four snapshots—are already taking some steps to use evaluation systems to support teacher growth. But additional actions must be taken by these as well as other states. Some actions can be through policy requirements that encourage accurate, frequent, actionable feedback, while others must be through the communication, support, and monitoring of these policies. While states' spheres of influence and capacity will vary, they can select and tailor high-impact actions for their specific context.

Policy Actions:

- **Ensure that teachers receive accurate observation ratings and meaningful feedback.** In order for evaluation systems to make good on the promise of improving the quality of teaching and learning, observers must correctly identify specific areas of strength and weakness and use that information to help teachers engage in professional growth. Three strategies for doing this are:

  - *Require high-quality observer training and certification assessments with a high bar for passage.* States should make rigorous *ongoing* observer training and certification mandatory for evaluation, as Tennessee has done, in order to ensure that all educators trust the feedback they receive, and the feedback meaningfully differentiates teacher performance to accurately inform their development. As in Tennessee, state training and certification should go beyond observer accuracy and also ensure capacity to effectively analyze evaluation data, provide targeted meaningful feedback, and design/recommend professional learning opportunities based on the results.

  - *Incorporate other content-area observers in the evaluation process.* Only four states with a policy requiring a connection between evaluation and development currently require multiple observers. Even in states that allow multiple observers, LEAs have not yet taken full advantage of the flexibility. Including additional observers has been shown to not only improve feedback accuracy, but also educators' perceptions of the evaluation process. States can forego hiring costly external observers by identifying and leveraging effective teacher leaders for evaluation so long as they pass rigorous certification protocols.

  - *Require or allow the use of student surveys.* Given that students observe instruction every day of the school year, their feedback can provide teachers and their instructional leaders with powerful information on their practice. While many states allow student surveys to be used in evaluation, few actually require them, and among those states that do allow their use, many LEAs have chosen not to include them. When thoughtfully designed and administered, student surveys can add formative value to the evaluation process while also increasing its accuracy and helping students feel empowered. States should employ developmentally-appropriate student surveys that have been deemed valid and reliable for the purpose of assessing teacher practice, as Massachusetts has done, and ensure teachers and instructional leaders are trained on how best to interpret and use the data for continuous improvement.

- **Ensure teachers receive frequent, timely feedback on their practice.** Feedback from evaluation systems can be made more impactful for development through the following actions:

  - *Require multiple observations annually for all teachers.* Increasing frequency does not have to mean increasing the burden on observers. Assuming a strong system for ensuring observer accuracy, states could allow flexibility so that consistently strong performers can waive some observations, customize their observations to focus on particular teaching practices, or only receive short rather than full-length observations. Additionally, observers could assess different components of educator practice in different observations rather than all at once, ideally based on areas that arose as priorities for improvement at the individual, school, and/or LEA-level in the prior year. This approach has the dual benefit of providing educators, particularly those new to the profession, with targeted feedback on a discrete and manageable set of skills to practice. Finally, allowing multiple observers, such as trained teacher leaders, can enable more frequent observations while lessening the burden on school principal observers.

  - *Expanding teacher leadership opportunities to encourage more informal observations.* States should incentivize and assist LEAs in creating formal teacher leadership pathways. Teacher leaders can provide their colleagues with more frequent informal or low-stakes feedback on their practice throughout the school year and lead professional learning to help implement new, more rigorous, academic standards. One way states can do this is by helping LEAs identify effective teacher leaders based on consistent demonstration of high performance, to lead teacher development, and encouraging school principals to leverage them further, as Louisiana and Tennessee have done.

- **Use formal evaluation system structures to drive a focus on development.** Several useful structures to aid teacher development already exist in many states' evaluation systems but are not being tapped to their full potential. Two ways to do this include:

  - *Require professional growth processes based on evaluation data for all teachers.* While many states require struggling teachers to be on a "professional growth plan," these are actually "professional intervention plans" and should be labeled accordingly, so that having a plan for professional growth is not synonymous with poor performance. Even in states that require growth plans for all teachers, school leaders may not implement these policies consistently or meaningfully. There are several reasons why this may occur, from a deficit of skills on how to effectively analyze evaluation data and plan appropriate professional learning to a lack of clarity about how to use the process effectively to support teacher development. To make these plans meaningful rather than compliance-driven, states should emphasize, provide guidance for, and monitor a growth *process*—as opposed to a *plan*—which should take place at the LEA, school, and individual teacher levels and allow for ongoing assessment and reflection, including setting and revisiting interim goals. At the individual level, teachers should lead the process with input from and ultimate approval by their observer. And professional growth planning should place student needs front and center.

  - *Ensure post-observation conferences shortly after observations occur.* Face-to-face debriefs that take place soon after the observation allow the observer and teacher to have the observation fresh in their memories, and make it more likely that the feedback discussed can be incorporated into practice in a timely manner. The structure also provides space for the observer and teacher to discuss and come to consensus around areas for growth and next steps, which is often more meaningful than simply receiving written feedback.

- **Balance objectivity and fairness in system design.** States have a responsibility to ensure that evaluation systems are as fair and objective as possible. These two aspects of evaluation can support each other, but can be at odds with each other as well. In striving for objectivity, some states have attempted to remove human judgment from the teacher evaluation equation, which may actually lead the systems to be less fair. States should take a lesson from Louisiana and allow for human discretion in evaluation and help ensure school leaders have sufficient skills and capacity to support and assess teachers' professional growth.

- **Enshrine in policy the vision of teacher evaluation as a development tool.** The evidence to date confirms that just having a state policy requiring that evaluation must inform teacher professional development is insufficient for ensuring this connection actually occurs. However, having a policy to point to when messaging the importance of evaluation for development can help signal that the state's evaluation system goals are aligned with educators' goals.

Communication Actions:

- **Clearly communicate evaluation as a tool for improvement.** States just beginning this work should communicate evaluation as a tool for growth and professional learning, as Colorado has done, rather than as an "I gotcha," before applying stakes to the results. States should then show teachers that this message is real by communicating the strategies the SEA is employing to ensure evaluation systems provide teachers with accurate ratings, actionable feedback, and support for engaging in meaningful professional learning opportunities connected to that feedback. But it is not too late for those states that have already come out strong on accountability to move toward a better balance. States must clearly communicate the measures and teaching indicators on which educators will be evaluated. While SEAs can tap local leaders to deliver these messages, SEAs should not rely on them to be the sole source of communications with teachers on this topic: teachers must know that these are statewide goals, and the SEA is committed to their attainment.

- **Engage in two-way dialogue with key stakeholders throughout implementation.** Obtaining input and feedback from stakeholders on an ongoing basis allows states to anticipate potential implementation issues and to make adjustments when they occur. Such stakeholders may include teacher and principal associations, school board members, and state legislators, in addition to educators themselves. States can work with stakeholders to develop key messages around educator evaluation as a tool for professional growth, as Colorado did during the early stages of implementing its system. They can also use stakeholder feedback to adjust communication strategies, as Kentucky did with its newsletters, or even make larger modifications to their systems, as Louisiana and Tennessee have done.

- **Coordinate messaging about evaluation across agency departments to ensure consistency.** The educator effectiveness/evaluation, professional development, and curriculum and assessment offices within state education agencies often operate in silos, despite their work being intricately connected at the school level. The result is that teachers often receive separate communications from each department highlighting disparate, disconnected priorities for what they should focus on, and potentially even different messaging about the same initiatives. Instead, the message that evaluation is a tool for feedback and continuous professional learning must be embedded in communications about student academic standards and the like. Bringing the various SEA divisions related to teacher practice and improvement together, as Louisiana and Kentucky have done, can help ensure that this occurs and will increase the staying power of educator evaluation systems, academic standards, and other initiatives SEAs are undertaking to improve teaching and learning. Even without this type of structure, SEAs' various divisions can and should coordinate their communications to clearly and coherently frame the goals and priorities of evaluation systems in the context of other work.

Support Actions:

- **Integrate teacher evaluation with other teaching and learning priorities, and support LEAs and schools in doing the same.** SEAs can ensure evaluation systems are viewed and used as development tools if they:

  - *Prioritize collaboration across all SEA divisions related to teaching.* Real collaboration across and integration of the SEA's various initiatives is key to improving teacher practice and student learning. One way states can do this is by aligning supports for evaluating teaching practices with supports for academic standards, as Louisiana has done. Such a strategy is likely to help drive the message that evaluation is for support, rather than just for accountability. Another way is to use evaluation system data to identify teacher leaders to deliver professional development trainings on state teaching and learning initiatives, as Tennessee and Louisiana have done.

  - *Help LEAs and schools understand how evaluation fits into the larger goal of improving teaching and learning.* States must adequately train school and LEA staff on the various standards they have adopted (teaching, professional learning, and academic) and how these initiatives can and should support each other, and then equip them with strategies and tools for making these connections like Louisiana has done by providing expanded content observation rubrics with key "look fors" in math and ELA.

- **Prioritize ongoing in-person support for observers.** States must build the skills of observers to accurately assess teacher practice, provide meaningful feedback, and connect the feedback to high-quality professional learning opportunities, and then follow up to monitor educator progress beyond the required initial training and certification. States should follow the lead of Delaware and North Carolina and offer ongoing trainings throughout the year so that observers maintain their skills. While these trainings would be optional for most observers, states could make them mandatory for those most in need of support. As part of that process, states could also invest in and support teacher leaders to provide educators with ongoing formative feedback and professional development connected with their evaluation results, as Louisiana has done.

- **Use data from evaluation systems to inform professional development, and help LEAs do the same.** Evaluation systems provide a wealth of data that can inform professional development at the individual, team, school, LEA, and state level. States with access to evaluation data at the individual teaching practice level should analyze them to identify trends in teacher strengths and weaknesses, and develop PD to address the areas where teachers were most in need of improvement, as New Jersey has done. Even states without direct access to evaluation data can provide technical assistance and trainings to LEAs and school leaders to perform these types of analysis. States should ensure observers and educators are made aware of—and trained on, if necessary—the professional learning resources and opportunities available to them.

- **Leverage technology to better integrate evaluation and development systems.** To help scale up efforts to connect evaluation and development, several states are using online platforms that "tag" professional learning resources to align with the indicators in the observation rubric. States should help ensure that educators are aware of and see the value of the resources made available to them. However, online resources should never replace in-person supports, such as coaching, particularly for higher-level teaching skills. Still, burgeoning research indicates that allowing video observations in addition to "live" observations can help teachers and observers develop a common frame of reference when discussing feedback in post-observation conferences, and could help build trust in the development aspect of the evaluation process.[105]

- **Ensure any development opportunities aligned to expected teaching practices are high-quality.** Technology has made it increasingly easy to share resources and opportunities for professional learning aligned to specific teaching practices in observation rubrics. However, states must vet any such resources prior to distributing them widely. Strong vetting can occur through empowering proven teacher leaders to co-create resources, having internal or external experts involved in their development, or instituting a rigorous process to source them from educators in its schools, as Kentucky has done.

- **Create structures and supports for team-based collaboration based on evaluation results.** Requiring professional growth plans and post-observation conferences are two ways in which states can help create structures for collaboration. But states can do more to help school leaders and educators work collaboratively toward school-wide, team, and individual goals as part of the evaluation process. For example, states could follow Tennessee and Colorado's lead and offer tools, guidance, and/or support to LEAs and school leaders on how to pair or group teachers based on evaluation results to create true "professional learning communities." As initial findings from Tennessee show, such structures can have an impact on teacher practice and student learning. And states should go further and provide guidance to LEAs around possibilities for re-envisioning the use of teachers' time, and consider providing incentives for LEAs to act on it.

- **Where possible, leverage external organizations and regional centers to maximize capacity and resources.** Despite good intentions, many SEAs are strapped for personnel and financial resources to do this important work. To overcome this challenge, states can partner with external organizations to assist with evaluation-related efforts, such as assessing teacher perceptions of evaluation efforts (as Tennessee did with SCORE and Vanderbilt), or providing in-person training sessions (as Colorado did with the Colorado Education Initiative). SEAs can also partner with regional education service centers or cooperatives to further their ability to support LEAs in growing teachers' practice, as Colorado has done, although they have to revisit how they manage their relationships with these entities to ensure their services are aligned to state priorities and efforts.

Monitoring Actions:

- **Institute processes to gauge whether teacher evaluation systems are being used to inform development.** After building observers' skills, SEAs should attempt to ensure that they are providing teachers with quality feedback and aligned learning opportunities. States can then use the information collected to determine where additional communication and support is needed. While these efforts will look different by state context, states could:

  - *Solicit feedback from teachers and school leaders on evaluation, and its connection to development.* States should regularly assess educators' perceptions of whether evaluation is being used as a tool for support, as Tennessee has done with its surveys, and then act on the results by targeting interventions and supports to those LEAs where teachers gave the poorest ratings. States could also follow Delaware's lead and interview a representative sample of schools within LEAs to dig deeper into how evaluation is being used to drive professional learning at the local level.

  - *Create tools or mechanisms to track or selectively audit observation data and feedback quality.* For example, Delaware developed a rubric to review observer feedback during monitoring visits, and states such as Tennessee and New York compare schools' observation and student growth scores to identify any misalignment between them. While Delaware can audit feedback more readily due to its small size, other states could audit a representative sample or assess feedback quality using educator surveys.

- **Use evaluation systems to assess quality and impact of specific professional development activities.** While some states are tracking educators' perception of PD quality or their participation in professional learning opportunities, few states are following up to assess impact of PD on teacher practice. But teacher evaluation systems could be a tool for measuring impact. In our interviews, Tennessee was the only state that mentioned tracking the impact of its training on new state standards using teacher evaluation data. But even if states do not have direct access to evaluation data, as Tennessee does, they can still help LEAs use their data to assess the quality and impact of educators' professional learning experiences over time.

- **Ensure school leaders are held accountable for teacher evaluation and development.** Many states are just beginning to fully roll out new principal evaluation systems. States should take the opportunity to develop policies and guidance to ensure these systems fully capture school leaders' teacher evaluation and professional learning responsibilities, as Tennessee has done. States should employ the lessons learned from evaluation implementation to help support school leaders in the areas in which they are accountable.

FINAL THOUGHTS

State education agencies are beginning to embrace the notion that both accountability and development play important roles in ensuring that evaluation systems have their intended effect of improving the quality of teaching for all students. While leading with development may make it easier to create the conditions for teacher growth to occur, even states that initially focused on accountability are finding ways to re-envision their evaluation systems to have a greater focus on support.

SEAs are also recognizing that while LEAs and schools will play the largest role in ensuring evaluation data are used to achieve the dual goals of accountability and development, those efforts will be more effective and more likely to happen at scale if states take on several roles as well: as a messenger, provider of support, and monitor of local efforts.

States need to be intentional in this work and consider the evidence on what types of professional development are most likely to truly grow teachers' knowledge and instructional quality. For example, while it is critical to move away from "one-size-fits-all" professional development, simply offering a platform connecting educators to online professional learning resources based on evaluation data will not be enough; this is only a move to "one-size-fits-most." Teachers must also have in-person support, including opportunities to practice what they have learned in a low-stakes environment, along with accurate, timely feedback on their progress. While ensuring individual teachers receive targeted feedback and learning opportunities is critical, these types of professional learning opportunities need not all be individualized. States should also analyze aggregate data to assess and plan to address professional learning at the state level, and assist and encourage LEAs and schools to do the same at their respective levels.

Doing this work well requires substantive commitment and capacity. Currently, many SEAs may feel they have more of the former than the latter particularly with Race to the Top's end.[106] And the recent reauthorization of the Elementary and Secondary Education Act—the Every Student Succeeds Act (ESSA)—no longer lists general provision of professional development for teachers as an allowable use of state funds.

However, states may use ESSA Title II funds for developing and improving evaluation and support systems that provide educators with "clear, timely, and useful feedback," or for assisting LEAs in doing so.[107] While this is good news for state education agencies, it does not mean they have everything they need to do this work well. State legislatures and governors, philanthropic organizations, technical assistance providers and others must support SEAs as they begin or bolster efforts to re-envision evaluation systems as tools for improvement and help generate a new narrative and culture around them: beyond ratings for personnel decisions and toward meaningful feedback for ongoing teacher—and student—growth.

Interviews Conducted

The following interviews were conducted as part of research for this paper:

- **Arizona**—Eric Brooks, Director, Effective Teachers and Leaders Unit, and Susan Poole, Education Program Specialist, Arizona Department of Education

- **Arkansas**—Ivy Pfeffer, Assistant Commissioner for Human Resources, Educator Effectiveness and Licensure, Arkansas Department of Education

- **Colorado** (site visit)—Katy Anthes, Executive Director of Educator Effectiveness, and Toby King, Director of Educator Effectiveness, Colorado Department of Education

- **Connecticut**—Shannon Marimon, Director for Educator Effectiveness and Professional Learning, Connecticut Department of Education

- **Delaware** (site visit)—Christopher Ruszkowski, Chief Officer of the Teacher and Leader Effectiveness Unit, and Eric Niebrzydowski, Deputy Officer of the Teacher and Leader Effectiveness Unit, Delaware Department of Education

- **District of Columbia**—Etai Mizrav, Manager, Education Policy and Compliance Division of Elementary, Secondary, and Specialized Education, Office of the State Superintendent of Education

- **Florida**—Eileen McDaniel, Chief, Bureau of Educator Recruitment, Development & Retention, Florida Department of Education

- **Georgia**—Cynthia Saxon, Associate Superintendent of Teacher and Leader Effectiveness, Georgia Department of Education

- **Hawaii**—Sean Arai, Personnel Director, Hawaii Department of Education

- **Kentucky**—Christine Meisberger, Manager, Teacher & Leader Effectiveness Branch, and Cindy Parker, former Division Director, Next Generation Professionals, Kentucky Department of Education

- **Louisiana** (site visit)—Rebecca Kockler, Assistant Superintendent of Academic Content and Melissa Mainiero, Deputy Director of Academic Content, Louisiana Department of Education

- **Maine**—No response

- **Massachusetts**—Regine Philippeaux-Pierre, former Professional Development Coordinator, and Matthew Holloway, Specialist, Massachusetts Department of Elementary and Secondary Education

- **Michigan**—Abigail Groff-Blaszak, Director of Educator Talent and Policy Coordination, Michigan Department of Education

- **Minnesota**—Tyler Livingston, Educator Evaluation Supervisor, and Renee Ringold, School Support Specialist, Minnesota Department of Education

- **Mississippi**—No response

- **New Jersey**—Carl Blanchard, Director, Office of Evaluation, New Jersey Department of Education

- **New Mexico**—Matthew Montano, Director, Educator Quality Division, New Mexico Department of Education

- **New York**—Alexander Trikalinos, Program Manager, Teacher and Leader Effectiveness, New York State Department of Education

- **North Carolina**—Robert Sox, Professional Development Coordinator, North Carolina Department of Public Instruction

- **North Dakota**—No response

- **Ohio**—Julia Simmerer, Senior Executive Director, Center for the Teaching Profession, Ohio Department of Education

- **Oregon**—Unable to schedule prior to publication

- **Rhode Island**—Sandra Forand, former Educator Quality Fellow, and Lauren Matlatch, Education Specialist, Rhode Island Department of Education

- **South Carolina**—Kristin Joannes, Director of Teacher Evaluation, South Carolina Department of Education

- **South Dakota**—No response

- **Tennessee**—Courtney Seiler, former Deputy Director of Evaluation, Nate Schwartz, Chief Research and Strategy Officer, Paul Fleming, Assistant Commissioner of Teachers and Leaders, Tennessee Department of Education

- **Utah**—Linda Alder, Coordinator of Educator Effectiveness, Utah Department of Education

- **Virginia**—Patty Pitts, Assistant Superintendent for Teacher Licensure, Virginia Department of Education

- **West Virginia**—Monica Beane, Executive Director, Office of Educator Effectiveness & Licensure, Trent Danowski, Manager of Educator Development, West Virginia Department of Education

- **Wyoming**—No response

Notes

1 Kati Haycock, *Good Teaching Matters a Lot: How Well-Qualified Teachers Can Close the Gap* (Washington, DC: Education Trust, June 1998), http://eric.ed.gov/?id=ED457260; Raj Chetty, John N. Friedman, and Jonah E. Rockoff, "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," Working Paper No. 17699 (National Bureau of Economic Research, December 2011), http://www.nber.org/papers/w17699; Steven G. Rivkin, Eric A. Hanushek, and John F. Kain, "Teachers, Schools, and Academic Achievement," *Econometrica* 73, no. 2 (March 2005): 417–458, http://www.econ.ucsb.edu/~jon/Econ230C/HanushekRivkin.pdf.

2 MET Project, *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains* (Seattle, WA: Bill & Melinda Gates Foundation, January 2012), http://eric.ed.gov/?id=ED540960.

3 Ibid.

4 Kathryn M. Doherty and Sandi Jacobs, *2015 State of the States: Evaluating Teaching, Leading, and Learning* (Washington, DC: National Council on Teacher Quality, November 2015), http://www.nctq.org/dmsView/StateofStates2015.

5 Stephen Sawchuk, "Teachers' Ratings Still High Despite New Measures," *Education Week*, February 5, 2013, http://www.edweek.org/ew/articles/2013/02/06/20evaluate_ep.h32.html; Linda Lyons, "Teachers Favor Common Core Standards, Not the Testing," Gallup, October 28, 2014, http://www.gallup.com/poll/178997/teachers-favor-common-core-standards-not-testing.aspx; William J. Bushaw and Valerie J. Calderon, "The 46th Annual PDK/Gallup Poll of the Public's Attitudes Toward the Public Schools: Part II," *Kappan Magazine* 96, no. 2 (October 2014): 49–59, http://pdkintl.org/noindex/PDKGallupPoll_Oct2014.pdf.

6 Linda Lyons, "Teachers Favor Common Core Standards, Not the Testing," Gallup, October 28, 2014, http://www.gallup.com/poll/178997/teachers-favor-common-core-standards-not-testing.aspx.

7 Title II, Part A, section 2101(c)(4)(B)(ii) of the Every Student Succeeds Act includes several activities as allowable uses of funds. One is "(ii) Developing, improving, or providing assistance to local educational agencies to support the design and implementation of teacher, principal, or other school leader evaluation and support systems that are based in part on evidence of student academic achievement, which may include student growth, and shall include multiple measures of educator performance and provide clear, timely, and useful feedback to teachers, principals, or other school leaders, such as by— '(I) developing and disseminating high-quality evaluation tools, such as classroom observation rubrics, and methods, including training and auditing, for ensuring inter-rater reliability of evaluation results;'" "(II) developing and providing training to principals, other school leaders, coaches, mentors, and evaluators on how to accurately differentiate performance, provide useful and timely feedback, and use evaluation results to inform decision making about professional development, improvement strategies, and personnel decisions;" and "(III) developing a system for auditing the quality of evaluation and support systems." p. 119, https://www.gpo.gov/fdsys/pkg/BILLS-114s1177enr/pdf/BILLS-114s1177enr.pdf.

8 Kathryn M. Doherty and Sandi Jacobs, *2015 State of the States: Evaluating Teaching, Leading, and Learning* (Washington, DC: National Council on Teacher Quality, November 2015), http://www.nctq.org/dmsView/StateofStates2015.

9 Emily Douglas, "6 Lessons on the Power of Teacher Leadership," *Education Week*: K-12 Talent Manager blog, January 15, 2015, http://blogs.edweek.org/topschooljobs/k-12_talent_manager/2015/01/power_of_teacher_leadership.html; Joey Garrison, "Metro Moves

Forward on Teacher Firings," The Tennessean, October 15, 2014, http://www.tennessean.com/story/news/education/2014/10/14/metro-moves-forward-teacher-firings/17284835/.

10 Delaware Department of Education, *Performance Matters: A Report on "Year Two" of the Revised DPAS-II Educator Evaluation System* (Delaware Department of Education: Teacher and Leader Effectiveness Unit, October 2014), http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/domain/271/present%20and%20reports/DPAS-II_Year_2_Report_2014.pdf; New York State Education Department, "2014 Preliminary Statewide Evaluation Results Released," press release, December 16, 2014, http://www.nysed.gov/news/2015/2014-preliminary-statewide-evaluation-results-released;http://www.michigan.gov/documents/mde/Educator_Effectiveness_Ratings_Policy_Brief_403184_7.pdf; Florida Department of Education, *Annual Legislative Report on Teacher Evaluation* (2013–14 School Year) (Florida Department of Education, 2014), http://www.fldoe.org/core/fileparse.php/7503/urlt/1314AnnualLegisReportTeacherEval.pdf.

11 Chad Alderman and Carolyn Chuong, *Teacher Evaluations in an Era of Rapid Change: From "Unsatisfactory" to "Needs Improvement"* (Washington, DC: Bellwether Education Partners, August 2014), http://bellwethereducation.org/sites/default/files/Bellwether_TeacherEval_Final_Web.pdf.

12 Morgaen L. Donaldson, *So Long, Lake Wobegon?: Using Teacher Evaluation to Raise Teacher Quality* (Washington, DC: Center for American Progress, June 2009), https://www.americanprogress.org/wp-content/uploads/issues/2009/06/pdf/teacher_evaluation.pdf; Matthew A. Kraft and Allison F. Gilmour, "Revisiting the Widget Effect: Teacher Evaluation Reforms and the Distribution of Teacher Effectiveness," working paper (Brown University, February 2016).

13 Education Counsel, *The Quality Framework: A Tool for Building Evaluation Systems that Improve Instruction* (Washington, DC, Education Counsel, April 2014), http://education.ky.gov/CommOfEd/fri/Documents/EducationCounsel%20Quality%20Framework%20Kentucky.pdf.

14 Alexander Russo, "Washington Post VAM-Based Teacher Firing Claim Might Cause Confusion," The Grade (blog), *Washington Monthly*, November 17, 2015, http://www.washingtonmonthly.com/the-grade/2015/11/washington_post_vambased_teach058662.php.

15 *Teachers Know Best: Teachers' Views on Professional Development* (Seattle, WA: Bill & Melinda Gates Foundation, 2015), http://collegeready.gatesfoundation.org/wp-content/uploads/2015/04/Gates-PDMarketResearch-Dec5.pdf.

16 Thomas Dee and James Wyckoff, "Incentives,

Selections, and Teacher Performance: Evidence from Impact," Working Paper 19529 (National Bureau of Economic Research, October 2013), http://curry.virginia.edu/uploads/resourceLibrary/16_Dee-Impact.pdf.

17 Melinda Adnot, Thomas Dee, Veronica Katz, and James Wyckoff, "Teacher Turnover, Teacher Quality, and Student Achievement," Working Paper 21922 (National Bureau of Economic Research, January 2016), http://www.nber.org/papers/w21922.

18 Eric S. Taylor and John H. Tyler, "Can Teacher Evaluation Improve Teaching?" *Education Next* 12 (Fall 2012), http://educationnext.org/can-teacher-evaluation-improve-teaching/.

19 Tennessee Department of Education, *Tennessee Educator Survey Report* (Tennessee Department of Education: Division of Data and Research, 2015), http://tn.gov/assets/entities/education/attachments/data_survey_report_2015.pdf.

20 Ibid.

21 Ibid.

22 Kathryn M. Doherty and Sandi Jacobs, *2015 State of the States: Evaluating Teaching, Leading, and Learning* (Washington, DC: National Council on Teacher Quality, November 2015), http://www.nctq.org/dmsView/StateofStates2015.

23 TNTP, *The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development* (New York: The New Teacher Project, 2015), http://tntp.org/assets/documents/TNTP-Mirage_2015.pdf; Eric S. Taylor and John H. Tyler, "Can Teacher Evaluation Improve Teaching?" *Education Next* 12 (Fall 2012), http://educationnext.org/can-teacher-evaluation-improve-teaching/.

24 MET Project, *Ensuring Fair and Reliable Measures of Effective Teaching* (Bill & Melinda Gates Foundation, January 2013), http://collegeready.gatesfoundation.org/wp-content/uploads/2015/05/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf; Grover J. Whitehurst, Matthew M. Chingos, and Katharine M. Lindquist, "Getting Classroom Observations Right," *Education Next* 15, no. 1 (Winter 2015), http://educationnext.org/getting-classroom-observations-right/.

25 MET Project, *Ensuring Fair and Reliable Measures of Effective Teaching* (Bill & Melinda Gates Foundation, January 2013), http://collegeready.gatesfoundation.org/wp-content/uploads/2015/05/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf; Heather Hill and Pam Grossman, "Learning from Teacher Observations: Challenges and Opportunities Posed by New Teacher Evaluation Systems," *Harvard Educational Review* 83, no. 2 (July 2013): 371–384, http://cepr.harvard.edu/publications/learning-teacher-observations-challenges-and-opportunities-posed-new-teacher;

Laura Bornfreund, *An Ocean of Unknowns: Risks and Opportunities in Using Student Achievement Data to Evaluate PreK–3rd Grade Teachers* (Washington, DC: New America, May 2013), http://fcd-us.org/sites/default/files/Ocean%20of%20Unknowns-%20Bornfreund.pdf.

26 Jess Wood, e-mail to Kaylan Connally, January 5, 2016.

27 MET Project, *Asking Students About Teaching: Student Perception Surveys and Their Implementation* (Bill & Melinda Gates Foundation, September 2012), http://www.metproject.org/downloads/Asking_Students_Practitioner_Brief.pdf.

28 MET Project, *Asking Students About Teaching: Student Perception Surveys and Their Implementation* (Bill & Melinda Gates Foundation, September 2012), http://www.metproject.org/downloads/Asking_Students_Practitioner_Brief.pdf; Tripod Education Partners, "Frequently Asked Questions," 2014, http://tripoded.com/general-faq/; Jeff Schulz, Gunjan Sud, and Becky Crowe, *Lessons from the Field: The Role of Student Surveys in Teacher Evaluation and Development* (Washington, DC: Bellwether Education Partners, May 2014), http://bellwethereducation.org/sites/default/files/Bellwether_StudentSurvey.pdf.

29 *Teachers Know Best: Teachers' Views on Professional Development*, (Seattle, WA: Bill & Melinda Gates Foundation, 2015), http://collegeready.gatesfoundation.org/wp-content/uploads/2015/04/Gates-PDMarketResearch-Dec5.pdf.

30 Interview with Regine Philippeaux-Pierre and Matthew Holloway, Massachusetts Department of Education, January 8, 2015; Massachusetts Department of Elementary and Secondary Education, Educator Effectiveness e-Newsletters (February 2013–October 2015), http://www.doe.mass.edu/edeval/communications/newsletter/.

31 Massachusetts Department of Elementary and Secondary Education, "Communications and Engagement Resources," http://www.doe.mass.edu/edeval/communications/.

32 Interview with Cindy Parker, Kentucky Department of Education, March 10, 2015; Interview with Christine Meisberger, Kentucky Department of Education, October 13, 2015.

33 Interview with Cindy Parker, Kentucky Department of Education, March 10, 2015.

34 Ibid.

35 Peter Youngs, *Using Teacher Evaluation Reform and Professional Development to Support Common Core Assessments* (Washington, DC: Center for American Progress, February 2013), https://www.americanprogress.org/wp-content/uploads/2013/01/YoungsCommonCore.pdf.

36 Massachusetts Department of Elementary and Secondary Education, "ESE Model Feedback Instruments & Administration Protocols," http://www.doe.mass.edu/edeval/feedback/surveys.html.

37 Massachusetts Department of Elementary and Secondary Education, "Educator Evaluation: Additional Resources and Archives," http://www.doe.mass.edu/edeval/resources/.

38 Interview with Eric Brooks and Susan Poole, Arizona Department of Education, January 28, 2015; Arizona Department of Education, "Presentation Resources from Summit IV," 2015, http://www.azed.gov/teacherprincipal-evaluation/summit-vi/.

39 Thomas R. Guskey and Kwang Suk Yoon, "What Works in Professional Development?" *Phi Delta Kappan* 90, no. 7 (March 2009): 495–500, http://www.k12.wa.us/Compensation/pubdocs/Guskey2009whatworks.pdf.

40 Interview with Ivy Pfeffer, Arkansas Department of Education, June 25, 2015; interview with Shannon Marimon, Connecticut Department of Education, January 21, 2015; interview with Eric Niebrzydowski, Delaware Department of Education, January 9, 2015.

41 Interview with Michele Purvis and Cynthia Saxon, Georgia Department of Education, January 9, 2015; Georgia Department of Education, "How Does Georgia FIP Align to the Teacher Assessment on Performance Standards (TAPS)?" http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/FIP%20Alignment%20to%20TKES.pdf.

42 Interview with Carl Blanchard, New Jersey Department of Education, November 15, 2015; New Jersey Department of Education, "Achievement Coaches Program," 2015, http://www.nj.gov/education/AchieveNJ/achievementcoaches/.

43 Interview with Eric Niebrzydowski, Delaware Department of Education, January 9, 2015.

44 Interview with Ivy Pfeffer, Arkansas Department of Education, June 25, 2015.

45 Interview with Sandra Forand, Lauren Matlach, and Melissa Denton, Rhode Island Department of Education, January 7, 2015.

46 Ibid.

47 Interview with Robert Sox, North Carolina Department of Public Instruction, April 30, 2015.

48 Interview with Tyler Livingston, Minnesota Department of Education, October 2, 2015.

49 Interviews with Christine Meisberger, Kentucky Department of Education, January 28 and October 13, 2015.

50 Tysza Gandha and Andy Baxter, *State Actions*

to Advance Teacher Evaluation (Atlanta, GA: Southern Regional Education Board, February 2016), http://publications.sreb.org/2016/160210_stateactionstoadvanceteachereval.pdf; Interview with Courtney Seiler, Tennessee Department of Education, January 7, 2015.

51 Kaitlin Pennington, *ESEA Waivers and Teacher-Evaluation Plans* (Washington, DC: Center for American Progress, May 2014), https://cdn.americanprogress.org/wp-content/uploads/2014/05/TeacherEvalWaivers-FINAL.pdf.

52 Interview with Alexander Trikalinos, New York Department of Education, January 16, 2015.

53 Interview with Michele Purvis and Cynthia Saxon, Georgia Department of Education, January 9, 2015.

54 Interview with Ivy Pfeffer, Arkansas Department of Education, June 25, 2015.

55 Interview with Michele Purvis and Cynthia Saxon, Georgia Department of Education, January 9, 2015.

56 New Teacher Center, "TELL Resource Library," http://teachingconditions.org/home/about-tell; Learning Forward Consulting, "Standards Assessment Inventory (SAI2) Information Packet," http://learningforward.org/docs/default-source/results/sai2-info-guide-revised.pdf?sfvrsn=0; Stephanie Hirsh, email to Melissa Tooley, August 3, 2015.

57 Andrew Vanacore, "Louisiana Launches Controversial New Teacher Evaluations to Reward Fear," *The Hechinger Report*, November 17, 2012, http://hechingerreport.org/louisiana-launches-controversial-new-teacher-evaluations-to-praise-fear/

58 Ibid.

59 Interview with Rebecca Kockler and Melissa Maniero, Louisiana Department of Education, March 20, 2015 , and interview with Rebecca Kockler, Louisiana Department of Education, May 29, 2015.

60 Ibid.

61 Louisiana Department of Education, "Observation and Feedback" from the Teacher Support Toolbox, http://www.louisianabelieves.com/resources/classroom-support-toolbox/teacher-support-toolbox/observation-feedback; Louisiana Department of Education, Louisiana Principals' Teaching and Learning Guidebook: A Path to High-Quality Instruction in Every Classroom, 2015, https://www.louisianabelieves.com/docs/default-source/teacher-toolbox-resources/2015-louisiana-principals'-teaching-learning-guidebook.pdf?sfvrsn=6.

62 Interview with Rebecca Kockler and Melissa Maniero, January 25, 2016.

63 Conversation with three teacher leader advisors,

Louisiana Department of Education, Teacher and Leader Summit, June 4, 2015.

64 Ibid.

65 Interview with Rebecca Kockler and Melissa Maniero, Louisiana Department of Education, March 20, 2015, and interview with Rebecca Kockler, Louisiana Department of Education, May 29, 2015.

66 John White, "Building and Supporting a System of Effective Teachers and Leaders," Louisiana Department of Education, https://www.louisianabelieves.com/docs/tap/brochure---louisiana-tap-brochure.pdf?sfvrsn=3; Joshua H. Barnett, Kellier C. Wills, and Peggy C. Kirby, *Comprehensive Educator Effectiveness Models that Work: Impact of the TAP System on Student Achievement in Louisiana* (Santa Monica, CA: National Institute for Excellence in Teaching, November 2014), http://www.niet.org/assets/PDFs/louisiana-tap-student-achievement.pdf?processed=1.

67 Colorado Department of Education, "Educator Effectiveness Key Messages," January 2015, http://www.cde.state.co.us/educatoreffectiveness/eekeymessages; Colorado Department of Education, "Educator Effectiveness: Communication Strategies and Pitfalls," http://www.cde.state.co.us/educatoreffectiveness/eecommunicationsstrategiesandpitfalls.

68 Jaclyn Zubrzycki, "Survey: Teacher Confidence in Evaluations Lower than Two Years Ago," *Chalkbeat Colorado*, May 7, 2015, http://co.chalkbeat.org/2015/05/07/survey-teacher-confidence-in-evaluations-lower-then-two-years-ago/#.Vm7yN-ODGko.

69 Ibid.

70 Interview with Katy Anthes, Colorado Department of Education, January 7, 2015.

71 Interview with Katy Anthes, Colorado Department of Education, October 7, 2015; School Improvement Network, "Edivate: The New PD 360," http://www.schoolimprovement.com/products/edivate/.

72 Interviews with Toby King, Colorado Department of Education, December 14, 2015 and October 26, 2015.

73 Andy Smarick and Juliet Squire, *The State Education Agency: At the Helm, Not the Oar* (Washington, DC: The Fordham Institute and Bellwether Education Partners, 2014), http://www.edexcellencemedia.net/publications/2014/State-Education-Agency-Helm-Not-Oar/State-Education-Agency-Helm-Not-Oar-FINAL.pdf.

74 U.S. General Accounting Office,"Education Finance: Extent of Federal Funding in State Education Agencies," October 14, 1994, http://www.gao.gov/assets/230/220301.pdf.

75 Colorado Educator Initiative, "Home Page," http://www.coloradoedinitiative.org/.

76 Interview with Katy Anthes, Colorado Department of Education, March 16, 2015.

77 Ibid.

78 Education First, *Giving Teachers the Feedback and Support They Deserve* (Washington, DC, Education First and Bill & Melinda Gates Foundation, May 2015), http://collegeready.gatesfoundation.org/wp-content/uploads/2015/05/EdFirst-Evaluation-Report.pdf.

79 State Collaborative on Reforming Education, "SCORE Releases Report Following Teacher Feedback Process," SCORE, June 11, 2012, http://tnscore.org/score-releases-report-following-teacher-evaluation-feedback-process/.

80 Tennessee Consortium, "First to the Top Surveys," 2011–2014, http://www.tnconsortium.org/projects-publications/first-to-top-survey/first-to-the-top-surveys/index.aspx.

81 Interview with Nate Schwartz, Tennessee Department of Education, March, 4, 2015.

82 Tennessee Department of Education, *Tennessee Educator Survey Report* (Tennessee Department of Education: Division of Data and Research, 2015), http://tn.gov/assets/entities/education/attachments/data_survey_report_2015.pdf.

83 Ibid.

84 Ibid.

85 Interview with Courtney Seiler, Tennessee Department of Education, January 7, 2015.

86 Ibid.

87 Interview with John Papay, Brown University, March 3, 2015.

88 John Papay, Eric S. Taylor, John Tyler, and Mary Lasky, "Learning Job Skills from Colleagues at Work: Evidence from a Field Experiment Using Teacher Performance Data," Working Paper 21986 (National Bureau of Economic Research, February 2016), http://www.nber.org/papers/w21986.

89 Tennessee Department of Education, "Is IPI Effective?" http://team-tn.org/ipi/is-ipi-effective/.

90 Tennessee Department of Education, "Instruction Partnership Initiative: Using Teacher Evaluation to Drive Instructional improvement," http://www.tncore.org/sites/www/Uploads/leadership_courses/ILC%20Class%201/KQ%205/Instructional%20Improvement.pdf.

91 Interview with Nate Schwartz, Tennessee Department of Education, March 4, 2015.

92 *Core Leadership: Teacher Leaders and Common Core Implementation in Tennessee* (Washington, DC: Aspen Institute, Education and Society Program, October 2014), http://www.leadingeducators.org/

resource/1413573107000/pubs_2014_oct/AES_LE_Core-Leadership-Tennessee_101614.pdf.

93 U.S. Department of Education, "Tennessee Trains 30,000 Teachers in the Common Core State Standards," Progress: Teachers, Leaders, and Students Transforming Education (blog), December 10, 2013, http://sites.ed.gov/progress/2013/12/tennessee-trains-more-than-30000-teachers-in-the-common-core-state-standards/.

94 Ibid.

95 Interview with Paul Fleming, Tennessee Department of Education, March 3, 2015.

96 Ibid.

97 Interview with Nate Schwartz, Tennessee Department of Education, March 4, 2015.

98 Interview with Eric Niebrzydowski, Delaware Department of Education, January 9, 2015.

99 Ibid.

100 Interview with Christopher Ruszkowski, Delaware Department of Education, March 17, 2015.

101 Ibid.

102 Linda Lyons, "Teachers Favor Common Core Standards, Not the Testing," Gallup, October 28, 2014, http://www.gallup.com/poll/178997/teachers-favor-common-core-standards-not-testing.aspx

103 Scott Sargrad, Samantha Batel, Melissa Lazarin, and Catherine Brown, *Investing in Educator Capacity: An Analysis of State Race to the Top Spending* (Washington, DC: The Center for American Progress, 2015), https://cdn.americanprogress.org/wp-content/uploads/2015/12/17114911/RTT-report.pdf.

104 U.S. Department of Education, "Awards," Race to the Top, March 2013, http://www2.ed.gov/programs/raceto-thetop/awards.html.

105 Thomas J. Kane, Hunter Gehlbach, Miriam Greenberg, David Quinn, and Daniel Thal, *The Best Foot Forward Project: Substituting Teacher Collected Video for In-Person Classroom Observations, First Year Implementation Report* (Cambridge, MA: Harvard University Center for Education Policy Research, 2015), http://cepr.harvard.edu/files/cepr/files/l4a_best_foot_forward_research_brief1.pdf?m=1443808234.

106 Alyson Klein, "Georgia Battles to Beat Race to Top Head Winds," *Education Week*, June 3, 2014, http://www.edweek.org/ew/articles/2014/06/04/33georgia_ep.h33.html.

107 Every Student Succeeds Act, Title II, Part A, section 2101(c)(4)(B)(ii), 119, https://www.gpo.gov/fdsys/pkg/BILLS-114s1177enr/pdf/BILLS-114s1177enr.pdf.