October 2011

State of the States

Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies

National Council on Teacher Quality

About NCTQ

The National Council on Teacher Quality is a non-partisan research and policy organization working to ensure that every child has an effective teacher.

Funding

Funding for this report was provided by the Bill and Melinda Gates Foundation and the Joyce Foundation.

For More Information

NCTQ is available to work with individual states to improve teacher policies. For more information, please contact Sandi Jacobs at sjacobs@nctq.org or 202-393-0020.

Executive Summary

State of the States: Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies

Each year, the National Council on Teacher Quality (NCTQ) publishes the *State Teacher Policy Yearbook*, a comprehensive examination of the state laws, rules and regulations that govern the teaching profession, measured against a realistic set of reform goals. For five years running, the full *Yearbook* compendium (www.nctq.org/stpy) presents the most detailed, thorough analysis of teacher effectiveness policy in the United States. In advance of the next *Yearbook*, to be released in January 2012, we offer a closer look at trends on teacher evaluation and effectiveness policies.

Across the states, there is unprecedented momentum towards developing and implementing teacher evaluation systems that factor student achievement into teacher ratings. While it is still too early to assess whether and to what extent states have actually been successful in developing and implementing meaningful performance-based teacher evaluation systems, in this report, NCTQ provides:

- a detailed picture of the teacher evaluation policy landscape across the states;
- an in-depth analysis of states with some of the most ambitious teacher effectiveness policies; and
- a set of early observations on the development and implementation of performance-based teacher evaluations.

The move to rethink how to evaluate teachers and explicitly tie assessments of teacher performance to student achievement marks an important shift in thinking about teacher quality. The change is significant because policymaking around improving teacher quality to date has focused almost exclusively on teachers' qualifications rather than on their effectiveness in the classroom and the results they get with students.

The landscape is changing. There are a host of policy recommendations focused on increasing the effectiveness of the teacher workforce that turn on the critical need to be able to evaluate and differentiate teacher performance reliably and consistently with clear criteria that include measures of how well teachers move students forward academically.

The federal Race to the Top (RTT) competition was certainly a first impetus for change in this area, but more than a few states have revised their policies on teacher evaluation without any federal incentives (and some RTT states haven't yet delivered). Regardless of motivation, the amount of evaluation reform activity is impressive:

In addition to providing a 50-state overview of teacher effectiveness policies, this paper looks more in-depth at the characteristics of the 17 states and the District of Columbia Public Schools[1] that are giving student achievement a significant, objective, meaningful and measurable role in how teacher performance is assessed:

Arizona, Colorado, Delaware, D.C., Florida, Idaho, Illinois, Indiana, Louisiana, Maryland, Michigan, Minnesota, Nevada, New York, Ohio, Oklahoma, Rhode Island and Tennessee

In these states, we examine:

- The frequency and timing of required evaluations
- State and district roles in and responsibilities for developing evaluation systems
- The specific measures included in evaluations and how significant a factor student achievement is in that mix
- The procedures for who conducts teacher evaluations and how they are conducted
- The required uses for teacher evaluation results

- Across the U.S., **32 states and the District of Columbia Public Schools (DCPS)[1] have made some change to their state teacher evaluation policy** in the last three years.
- Just two years ago, only 15 states required annual evaluations of all teachers, with some states permitting teachers to go five years or more between evaluations. As of this year, **24 states and DCPS require annual evaluations for all teachers**.
- Over this same short period of time, we've seen dramatic changes regarding the use of student achievement data to inform teacher evaluations. In 2009, 35 of the 50 states did not, even by the kindest of definitions, require teacher evaluations to include measures of student learning. Only four states could be said to be using student achievement as the preponderant criterion in how teacher performance was assessed. Today **23 states require that teacher evaluations include not just some attention to student learning**, but objective evidence of student learning in the form of student growth and/or value-added data.
- **Seventeen states and the DCPS** have adopted legislation or regulations that specifically require that student achievement and/or student growth will "significantly" inform or be the preponderant criterion in teacher evaluations.
- In **18 states and DCPS, teachers are eligible for dismissal based on teacher evaluation results**, although it is in only 13 of those states that teacher evaluations are explicitly tied to student performance.

Though the states we analyzed stand out for their specific focus on student achievement, it is still early to truly assess the state of the states on teacher evaluation. A few places, such as Delaware and D.C. Public Schools, are already implementing teacher evaluation systems. Others, such as Idaho and Minnesota, have just passed new requirements, and there has been no time for them to translate new policies into practice. Still other states, such as Colorado and New York, are deeply engaged in the process of developing evaluation instruments, negotiating specific system operating rules and in some cases, fighting hard battles to maintain commitment to a system where student learning is central to defining teacher performance.

While most policies are still very new, with many of the details to be determined, the changing landscape of teacher evaluation policy provides an opportunity to reflect on some of the early lessons:

- **Teacher effectiveness measures don't have to be perfect to be useful.**

  Are emerging teacher effectiveness measures perfect? No. But they are a marked improvement on evaluation systems that find 99 percent of teachers effective with little attention to a teacher's impact on students and offer little meaningful information on teachers' strengths, weaknesses and professional development needs.

1 Althought the District of Columbia has not enacted new state-level teacher evaluation policy, the District of Columbia Public School's evaluation system is among the most ambitious in the nation and thus is included here.

- **Insistence on comparability of measures for all teachers could cripple evaluation efforts.**
  The drive to identify or develop comparable measures for teachers regardless of grade or subject taught is understandable, but the more important emphasis ought to be on fair and valid measures.

- **Designing measures of student growth for non-tested grades and subjects is an important challenge facing states.**
  Thinking about the full complement of teachers – including K-2, social studies, special education, and non-core subject area teachers – states are approaching the challenge of how to develop fair and rigorous measures of student growth and achievement for all teachers in a variety of ways.

- **States shouldn't lose sight of the importance of classroom observations.**
  While there is a great deal of attention focused on linking value-added and student growth results to teacher evaluation, it is equally important to gather evidence observing behavior – what teachers do and what students are learning in the classroom – during classroom observation.

- **In addition to providing actionable feedback to all teachers, perhaps the most useful initial capacity of new evaluations will be to discern the most and least effective teachers.**
  The precision of growth and value-added data may not be at a very high level of sophistication but that doesn't mean they should be discounted.

- **Stakeholder input is important – but bold leadership is even more important.**
  Nothing about building a truly effective teaching force is going to come easy and the reality is that teacher reform is being met with unparalleled, vocal opposition. While it is critically important to have stakeholder voices represented, it must be balanced with real leadership and technical expertise where necessary.

- **State review and approval of district evaluations may not be an adequate approach to ensuring quality and rigor.**
  State approval sounds like a good idea in states that leave it to districts to design a performance-based teacher evaluation system. But it may not be realistic given state capacity. These states may do better to provide specific tools, models and detailed frameworks for conducting and scoring teacher evaluations. States that have left districts to their own devices without any oversight are even more worrisome. There is a good reason to be skeptical that all districts in such states will have the capacity and will to implement strong evaluation systems on their own.

- **States should start with annual evaluations for all teachers and modify for highly effective teachers once the system is fully operational.**
  Modifying an evaluation system to allow for less than full fledged annual evaluation may be sensible in some states, given issues of capacity, but states shouldn't start out that way.

- **States and districts should use third party evaluators when possible.**
  A third party evaluator can provide important feedback on the evaluation process and important checks for principals and other administrators.

- **A scarlet letter isn't appropriate teacher effectiveness policy.**
  Some think parental notification for students whose teachers received ineffective ratings is good accountability policy. But this humiliation tactic does a tremendous disservice to the teaching profession. Teachers with unacceptable levels of performance should be dismissed.

- **Teacher evaluation policy should reflect the purpose of helping all teachers improve, not just low-performers.**

  Many states are only explicit about tying professional development plans to evaluation results if the evaluation results are bad. Good evaluations with meaningful feedback should be useful to all teachers.

- **States should anticipate and address the anxieties a new evaluation system creates for teachers.**

  Teachers, not unlike most of us, are afraid of the unknown. States can do more to anticipate fears and diminish tensions over performance-based evaluations.

- **Escape clauses need to be shorn up and loopholes closed that may undermine new teacher evaluation systems.**

  Whether intentional or accidental, loopholes are already visible in some states' evaluation policies that can undermine their intended rigor. Without quick action to shore up these identified weaknesses, states may find themselves disappointed with the results they achieve and/or fighting unnecessary battles.

- **States need to get on top of policy plans for equitable distribution of effective teachers now.**

  Without some proactive planning, the exact opposite of more equitable distribution could occur when evaluation results are out and highly-effective teachers are identified.

- **States need to attend to potential bias with systematic checks of their evaluation system; states also need to maintain flexibility to make adjustments to the system as needed.**

  We are at the beginning of a new policy era about which there is still much to learn. In light of that, states should implement checks to ensure their evaluation systems are fair and reliable. Evaluation systems need to be flexible enough to take advantage of what we learn and be able to adjust.

What this policy review and early lessons suggest is that performance-based teacher evaluation must be approached in a measured, realistic and transparent way. Performance measures are not perfect and good teachers are not the product of formulas. Conducting teacher performance evaluations that focus on the results and the behaviors that matter most will move us toward a system that recognizes and encourages effective instruction and prepares and values highly-effective teachers.

The policy implications of an evaluation system that truly measures teacher effectiveness are profound. If done well, and if policymakers act on the results, the consequences could change much of what is now standard practice in the teaching profession by setting the foundation for better targeted policies for struggling teachers, higher standards for teacher preparation programs and fair but rigorous policies for replacing persistently ineffective teachers. Compensating teachers based on effectiveness could help attract and retain the best teachers in the profession. A system that cultivates effectiveness will also be crucial to other reform efforts, from implementing new Common Core State Standards and promoting educational equity, to turning around low-performing schools.

State of the States:

Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies

Background: Tracking State Teacher Effectiveness Policy

Each year, the National Council on Teacher Quality (NCTQ) publishes the *State Teacher Policy Yearbook*, a comprehensive examination of the state laws, rules and regulations that govern the teaching profession, measured against a realistic set of reform goals. For five years running, the full *Yearbook* compendium (www.nctq.org/stpy) presents the most detailed, thorough analysis of teacher policy in the United States, covering topics related to teacher preparation, licensure, evaluation, career advancement, tenure, compensation, pensions and dismissal policies.

The next full nationwide analysis of state teacher policy will be released in January 2012, presenting the legislative, regulatory and policy developments achieved in 2011. In advance of the next *Yearbook*, we offer a closer look into what is shaping up to be a very important education policy trend. States are revamping teacher evaluations so that they are significantly based on student achievement and tying that information to decisions of consequence about tenure, compensation, professional development and advancement.

With unprecedented momentum toward performance-based teacher evaluation across the states, the goals of this paper are to:

- Present a detailed picture of the teacher evaluation policy landscape across the states;
- Highlight noteworthy state policies; and
- Provide early observations on current challenges.

It is important to note that the development of teacher evaluation systems is a moving target, and we are in a period of rapid change. This analysis is focused on states with existing legislation, regulations and written policy, not states pondering changes or in the process of developing legislation. In this report, we examine state policies adopted as of September 2011. Any subsequent changes will be noted in NCTQ's forthcoming 2011 *Yearbook*.

It is still too early to assess whether and to what extent states have actually been successful in developing and implementing meaningful performance-based teacher evaluation systems. Indeed, only a few states (such as Rhode Island, Tennessee and Delaware) are implementing their new evaluation systems for the 2011-12 school year, and only the District of Columbia Public Schools (DCPS) has already applied consequences to teachers whose evaluations have shown either exceptional or very poor performance. Most states engaged in implementing new teacher effectiveness policies are still developing the details of their teacher evaluation systems. However, we believe that there are some lessons to be learned for those states already working on this issue as well as for those not yet tackling it.

Introduction

The move to rethink how to evaluate a teacher's performance and explicitly tie assessments of teacher performance to student achievement marks an important shift in thinking about teacher quality. The demand for "highly qualified" teachers is slowly but surely being replaced by a call for highly *effective* teachers.

The change is significant because policymaking around improving teacher quality to date has focused almost exclusively on a teacher's qualifications – teacher credentials, majors, degrees and licensing. Those criteria would be all well and good if they were associated with positive gains in student learning. Unfortunately, by and large, they are not.[1]

Yet the landscape is changing. Accountability for student learning and research confirming the strong impact teachers can have on student achievement have moved the field toward a decidedly *performance-based* focus on teacher quality. There are a host of policy recommendations for the effectiveness of the teacher workforce – including increasing pay so as to attract and retain talent, improving teacher preparation programs and removing consistently ineffective teachers. All of these policies turn on the critical need to be able to evaluate teacher performance reliably and consistently with clear criteria that include measures of how well teachers move students forward academically.

The naysayers argue that new trends in teacher evaluation are just the latest version of teacher-bashing, in which teachers are blamed for all that is wrong with education. In their view, these new evaluation systems employ punitive measures that not only will fail to improve teacher effectiveness, but that will lower the esteem of the teaching profession and demoralize teachers. They also argue that evaluating teachers based on student growth and achievement holds both good and bad teachers responsible for a set of outcomes that neither a good teacher nor a bad teacher can control.

Ultimately, we assert, the defense of the status quo is to argue that teachers do not make a difference, a stance that a solid body of evidence clearly refutes. Effective teachers matter a great deal and ineffective teachers may matter even more.[2] The refrain that judging teacher performance based on student

1 Boyd, Daniel; Grossman, Pamela, Lankford, Hamilton, Loeb, Susanna, and Wyckoff, James. 2006. "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement." *Education Finance and Policy*, 1(2); Jonah E. Rockoff, and Douglas O. Staiger. 2006. "What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City." Working Paper 12155. Cambridge, MA: National Bureau of Economic Research; Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). "Teachers, schools and academic achievement." *Econometrica*, 73(2), 417–458; Clotfelter, C., Ladd, H., and Vigdor, J., 2007. "How and why do teacher credentials matter for student achievement?" (Working Paper No. 12828). Cambridge, MA: National Bureau of Economic Research.

2 See Nye, B., Konstantopoulus, S., & Hedges, L.V. (2004). "How large are teacher effects?" *Educational Evaluation and Policy Analysis*, 26(3), 237-257; Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). "Teachers, schools, and academic achievement." *Econometrica*, 73(2), 417-458; Rockoff, J. (2004). "The impact of individual teachers on student achievement: Evidence from panel data." *American Economic Review*, 94(2), 247-252; Sanders, W.L., & Horn, S.P. (1998). "Research findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for educational evaluation and research." *Journal of Personnel Evaluation in Education*, 12(3), 247-256.

performance is unfair misses a central point about the capacity and the promise of value-added data. We can now account for many of the issues teachers can and cannot control when we evaluate performance. There are teachers that consistently achieve significant growth with the most disadvantaged students, while similar students make no progress with other teachers.

One of the greatest shortcomings of teacher performance appraisals has been school systems' *unwillingness and inability* to differentiate instructional competency. In *The Widget Effect*, The New Teacher Project quantified what was already known anecdotally – that teacher effectiveness "is not measured, recorded, or used to inform decision-making in any meaningful way" in most schools across the country. Their study across a set of twelve districts in four states found that less than one percent of teachers received unsatisfactory evaluation ratings.[3]

Much like schools' tendency to "teach to the middle," teachers have been evaluated to the middle, with evaluation tools neither designed nor implemented with an eye towards identifying the most talented educators or those who struggle. The reality is that there is huge variation in teacher performance (which is, in fact, true in any profession). But the disregard for performance in education has bred massive dysfunction and has disastrous consequences for the health of the teaching profession and for student achievement, especially for students most in need of effective teachers.

A comprehensive system for measuring, differentiating and acting on individual teacher performance data that is designed to advance the highest performers, develop the middle and deny tenure/dismiss the lowest, absent improvement, requires the following ten key elements:

1. **A data system** that generates growth or value-added data for teachers and a protocol for incorporating other objective student data for teachers without value-added data;
2. **Evidence of student learning as the preponderant criterion** of the evaluation instrument;
3. **Operating rules** for teacher evaluations that truly differentiate teacher performance (i.e., doing away with a system with only two possible ratings, such as satisfactory or unsatisfactory);
4. Teacher evaluation ratings based to a significant extent on **objective student data** (not limited to standardized test scores), including student growth or value-added data and data from sources such as formative assessments, progress in the curriculum and random sampling of student work, as well as **classroom observations** focused on a set of observable standards that gauge student learning;
5. To the extent possible, use of **trained third-party evaluators** to enhance and supplement the quality of feedback and support, but not to supplant a principal's important responsibility;
6. A **probationary (pre-tenure) period of sufficient length** in order to accumulate adequate data on performance on which to base decisions about teacher effectiveness;
7. A clearly articulated process for making **data-based tenure** decisions;

3 The New Teacher Project, 2009, "The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness" at http://widgeteffect.org/.

8. Specification of the obligations of the district and principal to provide **support structures for teachers identified as poorly performing** and a pre-established timeline for how long such support should last;

9. Streamlined mechanisms for **dismissing consistently poor performers** without stripping teachers' right of appeal by discarding lengthy legal proceedings and keeping all decisions in the hands of those with educational expertise;

10. A **comprehensive communications plan** to increase public awareness of this new system and the problems it means to solve.

A successful performance management system, with fair and reliable evaluations of teacher effectiveness as the centerpiece, is essential to the fundamental goal of all school reform: ensuring that all students have access to effective teachers and schools so that they achieve to their highest potential.

A Changing Landscape:

State of the States on Teacher Effectiveness Policies

The 2010 federal Race to the Top (RTT) competition spurred unprecedented action among the states to secure a share of $4 billion. A significant portion of the competition focused on state efforts to improve teacher and principal effectiveness based on performance. As a result, there are some promising and important new state laws and regulations on the books aimed at rethinking how teachers should be evaluated, compensated, promoted, granted tenure or dismissed based on their overall effectiveness in the classroom, including, in significant ways, teachers' impact on growth in student achievement.[4]

Several RTT winners are clearly at the forefront of efforts to develop and implement performance-based teacher evaluations. Delaware, Florida, Rhode Island, Tennessee, and D.C. Public Schools, for example, all require annual evaluations of all teachers and require that annual evaluations include objective evidence of student learning – not as an option, but as the *preponderant criterion* for assessing teacher effectiveness.

But it isn't just RTT states that are engaged in this work. States such as Colorado, Louisiana and Oklahoma are also on the same track – but without RTT funds. And, unfortunately, there are also RTT winners, such as Hawaii, with little or no legislative or regulatory changes to show for its promises regarding great teachers and leaders.

Race to the Top may have been a first impetus for change in this area, but more than a few states have revised their policies on teacher evaluation without any federal incentives. In 2011, states including Idaho, Indiana, Michigan, Minnesota and Nevada passed new teacher evaluation legislation without federal carrots or sticks. The changes are likely to keep coming. The U.S. Department of Education's just-announced September 2011 flexibility will allow states to apply for waivers of some of the specific requirements of No Child Left Behind in exchange for demonstrating, among other things, that they are employing teacher/principal evaluation and support systems focused on the quality of instruction and student results.[5]

While much of this paper will focus on states with the most ambitious evaluation plans, we thought it worthwhile to first look at the lay of the land across *all* states' teacher evaluation policies.

4 See http://www2.ed.gov/programs/racetothetop/index.html.

5 See http://www.ed.gov/esea/flexibility. States applying for NCLB waivers must "commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems."

Figure 1. National Overview of Teacher Evaluation Policies

There have been major advances in state policy on teacher effectiveness. In 17 states and DCPS, which this report will focus on in-depth, teacher evaluations are to be "significantly" informed by student achievement or student achievement and growth are to be the preponderant criterion in teacher evaluations. Not all of the 33 states that have made some kind of teacher evaluation policy change since 2009 have made dramatic improvements.

Highlighted states are the focus of "A Closer Look" section in this paper.
* States with Race to the Top grants from the U.S. Department of Education.

STATE	State has made policy changes related to teacher evaluations 2009-2011	State requires annual evaluations for all teachers	State requires that teacher evaluations include objective evidence of student learning	State specifies that teacher evaluations are to be "significantly" informed by student achievement/growth	State requires that student achievement/ growth is the preponderant criterion in teacher evaluations	State requires teachers to be eligible for dismissal based on evaluation results
Alabama	✓	✓				
Alaska						✓
Arizona	✓	✓	✓	✓		
Arkansas	✓		✓			✓
California						
Colorado	✓	✓	✓	✓	✓	✓
Connecticut	✓	✓	✓			
Delaware*	✓	✓[6]	✓	✓	✓	✓
D.C.*[7]	✓	✓	✓	✓	✓	✓
Florida*	✓	✓	✓	✓	✓	✓
Georgia*	✓	✓				
Hawaii*						✓
Idaho	✓	✓	✓	✓	✓	
Illinois	✓		✓	✓		✓
Indiana	✓	✓	✓	✓		✓
Iowa						
Kansas						
Kentucky						
Louisiana	✓	✓	✓	✓	✓	✓
Maine	✓					
Maryland*	✓		✓	✓	✓	
Massachusetts*	✓		✓			
Michigan	✓		✓	✓	✓	✓
Minnesota	✓	✓[8]	✓	✓		
Mississippi						

6 While Delaware does not require a full-fledged summative evaluation every year (instead, every other year), the state does track whether teachers meet student growth expectations each year. If a teacher does not meet his/her growth expectations in a year when a summative evaluation would not normally be conducted, the failure to meet annual growth requirements triggers a full-fledged evaluation.

7 In NCTQ's annual *State Teacher Policy Yearbook*, NCTQ examines the District of Columbia's statewide policies under the responsibility of the Office of the State Superintendent of Education (see http://osse.dc.gov/seo/site/). At the state level, D.C. does not have teacher evaluation policies that meet the above criteria. However, because the District of Columbia Public Schools (DCPS) has significant teacher evaluation policies in place, we included DCPS in this analysis. When we reference D.C. in this paper, we are referring to DCPS.

8 While Minnesota describes an *annual* evaluation process for teachers, the state's policy establishes a three-year professional review cycle for each teacher that includes an individual growth and development plan, a peer review process, the opportunity to participate in a professional learning community, and at least one summative evaluation performed by a qualified and trained evaluator such as a school administrator. For the years when a tenured teacher is not evaluated by a qualified and trained evaluator, the teacher must be evaluated through a peer review. As written, it isn't clear that what will occur in the out years will be an adequate review of teacher performance.

STATE	State has made policy changes related to teacher evaluations 2009-2011	State requires annual evaluations for all teachers	State requires that teacher evaluations include objective evidence of student learning	State specifies that teacher evaluations are to be "significantly" informed by student achievement/growth	State requires that student achievement/ growth is the preponderant criterion in teacher evaluations	State requires teachers to be eligible for dismissal based on evaluation results
Missouri						
Montana						
Nebraska						
Nevada	✓	✓	✓	✓	✓	✓
New Hampshire	✓					
New Jersey		✓				
New Mexico						
New York*	✓	✓	✓	✓		✓
North Carolina*	✓	✓	✓			
North Dakota		✓				
Ohio*[9]	✓	✓	✓	✓	✓	
Oklahoma	✓	✓	✓	✓	✓	✓
Oregon	✓					
Pennsylvania		✓				✓
Rhode Island*	✓	✓	✓	✓	✓	✓
South Carolina						
South Dakota	✓					
Tennessee*	✓	✓	✓	✓	✓	✓
Texas						✓
Utah	✓	✓	✓			
Vermont						
Virginia	✓					
Washington	✓	✓				✓
West Virginia						
Wisconsin	✓					
Wyoming	✓	✓	✓			
TOTALS	**33**	**25**	**24**	**18**	**13**	**19**

9 Due to the pending referendum on its education reform bill, Ohio currently has two versions of its education code pertaining to teacher evaluation on its books. One version addresses removal of poorly performing teachers, but the other does not.

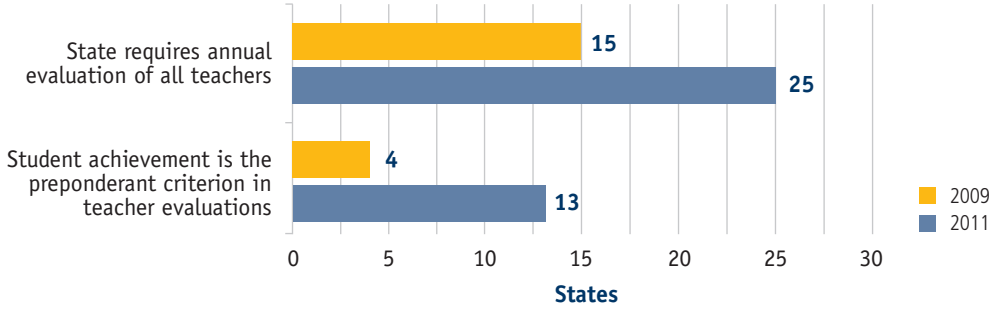
The data in Figure 1 indicate more than just a policy trend. The policy shift on teacher evaluations across the states since 2009 is dramatic. Just two years ago, only 15 states required annual evaluations of all teachers, with some states permitting teachers to go five years or more between evaluations.

Requiring annual formal teacher evaluations is one thing; making sure evaluations actually measure teacher effectiveness is another. Over this same short period of time, we've seen dramatic changes regarding the use of student achievement data to inform teacher evaluations. In 2009, 35 states did not, even by the kindest of definitions,[10] require teacher evaluations to include measures of student learning. In that same year, only four states could be said to be using student achievement as the preponderant criterion in how teacher performance was assessed, again, even with a loose and generous interpretation.[11]

In 2011, we see quite a different landscape. Twenty-three states and DCPS require that teacher evaluations include not just some attention to student learning, but objective evidence of student learning in the form of student growth and/or value-added data.

Figure 2. Shifting State Teacher Evaluation Policy 2009-2011

In 2011, 24 states as well as the District of Columbia Public Schools (DCPS) require annual evaluation of all teachers. Twelve states and District of Columbia Public Schools require that objective measures of student achievement be the preponderant criterion for assessing teacher performance in teacher evaluations.



Source: NCTQ 2009 *State Teacher Policy Yearbook*, research for 2011 *State Teacher Policy Yearbook*

10 In the 2009 *State Teacher Policy Yearbook*, NCTQ acknowledged that states included evidence of student data for any requirements that attempted to consider student performance. With the significant advances states have made in their requirements for measuring student performance, beginning in 2011, NCTQ will raise the bar on states in the *Yearbook*. States will receive credit for including objective evidence only if student data includes (but need not be limited to) student growth and/or value-added data (where applicable).

11 See NCTQ, 2009 *State Teacher Policy Yearbook* at www.nctq.org/stpy2009. The next comprehensive review of state teacher policies by NCTQ is due out in January 2012.

A Closer Look

at States with Ambitious Teacher Evaluation Policies

The highlighted states in Figure 1 have proposed some of the most significant changes in teacher evaluation policy in the nation and merit a more detailed examination: Arizona, Colorado, Delaware, D.C.,[12] Florida, Idaho, Illinois, Indiana, Louisiana, Maryland, Michigan, Minnesota, Nevada, New York, Ohio, Oklahoma, Rhode Island and Tennessee.

Each of these states has adopted legislation or regulations that specifically require teacher evaluations to include *objective evidence of student learning* and mandate that *student achievement and/or student growth will "significantly" inform or be the preponderant criterion* in teacher evaluations. The legislative or regulatory language leaves little room for misinterpretation.

Why are other states not included in this list? In particular, why are several RTT states not included?

- **We did not include states where evaluation projects are not statewide.** For instance, **Georgia** is not included because the evaluation system the state is implementing as part of its RTT grant is, at least at this point, limited only to those 26 districts (out of 181 total districts in the state) participating in RTT.
- **We did not include states where evaluation projects are only part of RTT proposals.** In **Hawaii**, the statewide teacher evaluation redesign proposed as part of its RTT application hasn't materialized in any significant way.
- **We also did not include states that left too much detail up for grabs on the kinds of objective measures to be included in teacher evaluations.** These states might be explicit in requiring the use of student achievement measures in teacher evaluations but were less explicit in requiring consistent use of objective measures or were unclear about how significant a factor student achievement will be in teacher evaluations – leaving what some would call significant discretion, but we could call loopholes, in the developing systems.

17 states and DCPS[12] plan to give student achievement a significant, objective, meaningful and measurable role in how teacher performance is assessed:

Arizona, Colorado, Delaware, D.C., Florida, Idaho, Illinois, Indiana, Louisiana, Maryland, Michigan, Minnesota, Nevada, New York, Ohio, Oklahoma, Rhode Island and Tennessee

12 Although the District of Columbia has not enacted new state-level policy teacher evaluation policy, the District of Columbia Public School's evaluation system is among the most ambitious in the nation and thus is included here.

- For example, as of July 2011, **North Carolina** has added a standard to its evaluation system requiring teachers to contribute to the academic success of students. But there is not much clarity about how current evaluation instruments ensure that "the work of the teacher results in acceptable, measurable progress for students based on established performance expectations using appropriate data to demonstrate growth." Furthermore, the new standard, unlike the state's other five performance standards, does not have a required performance rating associated with it.[13]

- Likewise, we didn't include **Massachusetts**. While an important RTT state grantee, Massachusetts's new regulations stating that "multiple measures of student learning, growth and achievement" must be one category of evidence in teacher evaluations leave too many details and too much discretion to individual evaluators to choose student achievement measures and make decisions about the adequacy of growth attained by individual teachers. Notably, the state's final regulations (unlike the drafts leading up to the final) removed language requiring that the student performance measures be a "significant" factor in teacher evaluations.[14]

- **Arkansas** began requiring student achievement and growth measures to be included in evaluations this year. While the state notes that the rules of the system shall "recognize" that evidence of student growth is "significant," the state allows individual teachers a good deal of discretion in choosing "artifacts" that will be used in an evaluation. Even though the state specifies that 50 percent of the artifacts must be "external assessment measures," it appears that even in grades and subjects for which state assessment data are available, these test results could be just an option for inclusion in the evaluation of a teacher. Telling may be the state's removal of language from the final bill that would have required 50 percent of the weight of teacher evaluations in Arkansas to be based on student test results.[15]

Though the states we are taking a closer look at stand out for their specific focus on student achievement, they may not necessarily be the farthest along on implementing such systems. Delaware and DCPS are already implementing teacher evaluation systems. Other states, such as Idaho and Minnesota, have just passed these new requirements, and there has been no time for them to translate new policies into practice. Still others, including Colorado and New York, are deeply engaged in the process of developing evaluation frameworks or instruments, negotiating specific system operating rules and, in some cases, fighting hard battles to maintain commitment to a system where student learning is central to defining teacher performance.

13 See http://sbepolicy.dpi.state.nc.us/.

14 See the draft regulations (http://aftma.net/wp-content/uploads/2010/12/Regs-on-Evaluation-of-Educators_version-1_April-16.pdf) in Massachusetts, which noted that "Student Performance Measures shall be a significant factor in the summative evaluation," compared with the final regulations at http://www.doe.mass.edu/lawsregs/603cmr35.html.

15 In explaining why the Arkansas Education Association (AEA) decided to support Arkansas' teacher evaluation bill, AEA notes that references to the weight of a teacher evaluation based on student test scores was removed from the bill, replaced with a provision that one-half of the "artifacts" considered in a teacher's evaluation shall be student test scores, which are "listed in only 2 of the 22 components of the teacher evaluation framework." Other "artifacts" include, among other options, self-directed research; contributions to parent, community or professional meetings; and participation in professional development.

Key Issues Related to Teacher Evaluation and Effectiveness Policy

1. What is the frequency and timing of required evaluations?
2. Who is responsible for development of the evaluation system (state or districts; if districts, is there a review and approval process)?
3. What measures are required?
4. How is student growth factored?
5. Is student achievement the preponderant criterion?
6. How many and what are the evaluation categories/ratings?
7. Who is responsible for conducting evaluations?
8. What observation procedures are specified (number and length, feedback conferences with teachers)?
9. What are the required uses for evaluation results (improvement plans, professional development, assistance, more frequent follow up evaluations for teachers with ineffective ratings)?
10. What employment consequences are tied to evaluation results (salary increases, bonuses, tenure decisions, licensing, dismissal policies)?

Frequency and Timing of Teacher Evaluations

An important aspect of developing policies to drive improvements in teacher effectiveness policy are systems that provide teachers with regular, actionable feedback for their own growth and development and help schools make meaningful, informed decisions about the performance of teachers.

First, states need to set an appropriate length to teacher probationary periods (at least four years). The timing of non-tenured teacher evaluations needs to be stated explicitly, and early observations of new teachers should be required so that teachers who are struggling or have specific training needs can receive support immediately.

Across the 17 states and DCPS, all but three – Illinois, Maryland and Michigan – require districts to conduct teacher evaluations annually for *all* teachers, with some modifications, some stronger than others. Though not optimal, some modifications to evaluations for highly effective veteran teachers may make good sense, provided that a state is annually collecting high-quality objective data on student and teacher performance and provided that a state never abdicates its right to evaluate any teacher in a given year no matter what the record of previous performance might have been.

- In **Delaware**, experienced teachers who earn a rating of "highly effective" on their most recent summative evaluation must receive a minimum of one announced observation each year but with a summative evaluation at least once every two years. However, the "student improvement" component must be evaluated every year, and teachers cannot be rated "effective" unless they have met growth targets annually. Furthermore, if a highly effective teacher does not achieve a satisfactory rating on the student improvement component, the teacher must receive a summative evaluation the following year.

- **Michigan** makes exceptions on the evaluation timeline for demonstrated highly effective teachers. A district can choose to evaluate a highly effective teacher once every other year as long as the teacher remains highly effective.

- In **Illinois**, non-probationary teachers are required to be evaluated once every two years; however, a "needs improvement" or "unsatisfactory" evaluation rating in any given year triggers a follow-up evaluation in the year following such a rating.

With a rigorous and well-implemented teacher evaluation system, some of these staggered timelines may be sensible. But states need to be careful not to go in this direction prematurely. Other states' requirements make less sense.

- In **Maryland**, teachers who have earned advanced professional certificates are required to be evaluated twice during the 5-year period of the certificate. Besides the fact that this policy stretches the time between evaluations to as many as 4 years, the criteria for earning an advanced professional certificate are not performance based. Advanced certificates in Maryland are awarded to teachers who earn an advanced degree, which research has shown time and again to have no correlation to student achievement, and which may say little about a teacher's effectiveness in the classroom.[16]

We see no basis for a state like Maryland to assume that teachers who have earned advanced certificates are effective, and as a general rule evaluate such teachers less frequently.

However, in most of these states, teachers undergo more scrutiny during the probationary period, particularly in their first years of teaching, before reaching tenure or permanent status.[17] This typically involves more frequent classroom observations, feedback conferences and written performance reports. Ideally, such policies would be specifically timed to help ensure that evaluations happen when they can be most useful as a real-time feedback opportunity.

Multiple observations is good policy. But multiple evaluations – the compilation of data from multiple sources to produce a rating – within the school year in an evaluation system that requires student achievement and growth data to be a critical factor is not realistic; the relevant data measures would neither be available nor meaningful early in the school year. Going forward, there may be a need to clarify nomenclature – as an evaluation and an observation are no longer synonymous in performance-based systems – and/or recalibrate the timing of some requirements. In addition, there may also be a need for states to better clarify the rules, specifying whether teachers must be observed multiple times for a summative evaluation or whether districts are expected to conduct multiple evaluations each year.

16 Clotfelter, C., Ladd, H., and Vigdor, J., 2007. "How and why do teacher credentials matter for student achievement?" (Working Paper No. 12828). Cambridge, MA: National Bureau of Economic Research; Rice, J. K. (2003, August). "Teacher quality: Understanding the effectiveness of teacher attributes." Washington, DC: Economic Policy Institute. Rivkin, S., Hanushek, E., and Kain, J. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica*, 73(2): 417-458.

17 While not all states use the term "tenure," the end of a teacher's probationary period generally has the same implications regardless of nomenclature.

Figure 3. Requirements and Timing of Evaluations for New Teachers

Nine of the states we examined specify deadlines for when evaluations must occur during the school year, and most require multiple evaluations or multiple observations of new teachers.

STATE	Number of annual evaluations or observations[18] required for probationary teachers	State specifies timing of evaluations for probationary teachers	Other requirements for probationary teacher evaluations
Arizona	2 evaluations	No	None specified
Colorado	2 documented observations and 1 evaluation	No	Written evaluation report
Delaware	Minimum of 2 announced observations and 1 unannounced observation resulting in a summative evaluation	No	Each announced observation includes a conference
D.C.	5 observations	First observation by administrator must occur between September 21 and December 1; first observation by master educator must occur between September 21 and February 1	A meeting between observer and teacher within 15 days of each observation
Florida	2 observations and evaluations in the first year of teaching	No	None specified
Idaho	1 evaluation with two parts (the first part includes input from parents and guardians of students as a factor and the second portion must be based on objective measures of student growth)	First part of evaluation must occur before January 1	None specified
Illinois	1 evaluation	No	None specified
Indiana	1 evaluation	Formal evaluation is required before December 31	If requested by teacher, additional evaluation may be scheduled on or before March 1
Louisiana	1 evaluation	No	Pre and post observation conference
Maryland	1 evaluation per semester	No	Each evaluation includes a conference
Michigan	1 evaluation (unless rated effective or highly effective for two consecutive year-end evaluations, including for non-tenured teachers) based on multiple observations	No	Mid-year progress report required for first year teachers
Minnesota	3 evaluations	The first evaluation must occur within 90 days of the beginning of teaching services	Not specified
Nevada	3 evaluations	Evaluations must be concluded not later than December 1, February 1, and April 1	Each evaluation includes a conference and written evaluation
New York	1 evaluation	No	Not specified
Ohio	2 evaluations	The first evaluation must be completed by January 15 with a written report to the teacher by January 25; the second must be scheduled between February 10 and April 1, with a written report to the teacher by April 10	Written report
Oklahoma	At least 2 evaluations	Evaluation required once prior to November 15 and once prior to February 10 of each year	Not specified
Rhode Island	At least 4 observations annually (combination of longer announced and shorter unannounced visits)	Conferences must be timed to be at the beginning of the school year, midway through the year and at the end of the year. One observation is required during the first semester prior to the mid-year conference	Three conferences are required as part of annual evaluation
Tennessee	6 observations	3 observations per semester	Post-observation conference required after each observation

18 This table includes language as stated in state law and regulations, which sometimes refers to multiple "evaluations" per year and sometimes multiple "observations" per year intended to *inform* a summative annual evaluation. It is unclear in some cases whether states are using the terms "observations" and "evaluations" interchangeably, or if some states are actually requiring multiple full-fledged summative evaluations per year (an unrealistic goal).

Roles and Responsibilities in Evaluation Design

Until this new wave of reform, states typically have specified little detail about teacher evaluations and few, if any, states have scrutinized their districts' various teacher evaluation systems. But state policymakers and state education officials are increasingly asserting themselves more into the specifics. Given the complexity and scope of evaluation systems, at a minimum, states ought to provide a model that districts can adopt and/or adapt, rather than leaving districts to fend for themselves.

Looking across states that have proposed the most ambitious performance-based teacher evaluation systems, states have adopted a diverse set of approaches to this balance between state and local interests. Most states still leave much of the details to districts, but several offer statewide systems or state-developed models that shape the evaluation criteria meant to keep student achievement and growth front and center in teacher evaluations.

Figure 4. State Models for Teacher Evaluation Design

There is a great deal of variation in how states are approaching the design of teacher evaluation systems, clearly illustrating that states have real options. There is nothing close to a one-size-fits-all teacher evaluation design, probably a good thing. It remains to be seen if there is an approach that works best.

STATE	Single, statewide teacher evaluation system	Presumptive state designed teacher evaluation model	State designed teacher evaluation model with district opt-in	District designed systems consistent with state framework and including state review and approval process	District designed system consistent with state framework
Arizona					✓
Colorado			✓		
Delaware	✓				
D.C.[19]					
Florida				✓	
Idaho				✓	
Illinois			✓		
Indiana			✓		
Louisiana	✓				
Maryland				✓	
Michigan		✓[20]			
Minnesota					✓
Nevada					✓
New York					✓
Ohio					✓
Oklahoma			✓[21]		
Rhode Island		✓			
Tennessee		✓			
TOTALS	**2**	**3**	**4**	**3**	**5**

19 As noted, this paper focuses on the District of Columbia Public Schools, not state-level policy. The charter school districts in the District of Columbia are not obligated to use DCPS' IMPACT system. DCPS is not included in the counts on this table as its evaluation model is not "statewide."

20 Michigan's law requires districts to implement an evaluation tool developed by the state. The law also states that if a school district has a local evaluation tool for teachers consistent with the state evaluation tool, the district can use that local evaluation tool. It is unclear how or whether a district would need to demonstrate that its tool is consistent.

21 Preliminary recommendations posted by the state of Oklahoma in September 2011 for public comment propose that the state name a default framework to serve as the qualitative assessment component that must comprise 50% of the total evaluation criteria required.

Delaware and Louisiana have developed specific evaluation processes and instrumentation that their districts are required to adopt. **Delaware**'s Performance Appraisal System (DPAS II) gives its districts all of the tools it will need, including evaluation guides, appraisal components and appraisal criteria for all evaluations. **Louisiana** is also designing its own statewide teacher evaluation system, known as COMPASS. Although not technically a statewide system, **D.C.**'s IMPACT system covers the more than 6,500 school-based employees in D.C. Public Schools.

Rhode Island has designed a model system for its districts that is intended to show districts how to meet state educator evaluation standards, essentially a *presumptive model* for teacher evaluation. The state has a full-fledged model developed, with processes and instrumentation that can be, and indeed are expected to be, adopted statewide. However, districts are not prohibited from putting forward alternative designs for approval, as long as they meet the same standards as the state model.

Tennessee is implementing its Tennessee Educator Acceleration Model (TEAM), which includes the TAP framework observation protocol, for which the state is providing training and certification to evaluators statewide. For the achievement and growth portion of the evaluation, teachers are evaluated using the Tennessee Value Added Assessment System (TVAAS). However, the state will allow districts to propose their own observation protocol with state review and approval.

Michigan's teacher evaluation law requires the state to develop a model teacher evaluation tool, although it remains unclear whether districts must demonstrate they have a comparable local instrument if they elect not to use the state's tool.

Three states (**Colorado**, **Illinois** and **Indiana**) can best be categorized as *opt-in* states. Districts are free to choose to adopt state evaluation models in lieu of designing their own comparable systems. The relevant distinction here is that in these states, the primary responsibility for formulating an evaluation consistent with state policy lies with the district. The state model is available as a resource (or as a default for districts that fail to act on their own to create their own system), *but not as a requirement*.

In eight of our ambitious states, individual districts bear the primary responsibility for designing local teacher evaluation systems, but based on the basic requirements and parameters laid out in state legislation and/or regulations. Frameworks for evaluation in these states vary in their levels of detail and prescriptiveness, but what is clear is that the state has no specific obligation to design and disseminate a full-fledged model for districts. While it remains to be seen whether some of these states decide that a state-developed model is a worthwhile investment, these states are providing various levels of guidance and technical assistance to districts.

Nevada requires each local board to "develop a policy for objective evaluations in narrative form" consistent with teacher evaluation law requirements around how often, how long, when during the school year, with what ratings (highly effective, effective, minimally effective or ineffective) and with what required feedback evaluations must occur. **Arizona** is providing a repository of evaluation instruments that comply with the state's evaluation framework, but they are not specifically required.

In three of the eight states where primary responsibility falls to districts to develop evaluation systems, **Florida**, **Maryland** and **Idaho**, districts are required to submit their evaluation systems to the state for review and approval. Maryland also provides a default approach to measuring professional and student growth should a district and collective bargaining unit fail to agree to a performance standard.

State Formulas for Performance-Based Evaluations

What sets these emerging state policies apart from anything the field has seen before is that states are not only rethinking their standards for effective teaching, but they are defining the specific measures of effectiveness and the respective values of those measures within a comprehensive performance appraisal.

At the center of new teacher evaluations is a mix of elements that the state requires to be included when assessing teacher effectiveness and for which the state has established specific values and weights. While many of the details of these policies are still very much under development in some states a set of *formulas*, of sorts, for teacher evaluation is emerging across the states.

Whether framed around the data captured by a mix of evaluation measurement tools (observations, surveys, growth scores etc.) or framed around topical components (instruction, classroom management, student achievement), these formulas begin to describe the operating rules of the systems and just how and to what extent student achievement and growth count towards overall assignment of effectiveness ratings to individual teachers.

Some states are quite prescriptive in spelling these formulas out in state statute or regulations.

- For example, **Tennessee** requires the following mix of measures:
  - A total of 50 percent of a teacher's annual evaluation must be based on student achievement data, of which 35 percent must rely on student growth data from the Tennessee Value-Added Assessment System (TVAAS). The remaining 15 percent must be based on other measures of student achievement (selected from menu of options adopted by the state).
  - 50 percent of a teacher's evaluation must be based on other criteria (observations, surveys, review of prior evaluations, conferences).
- **New York** requires a formula that weights key components differently:
  - A total of 20 percent of a teacher's annual evaluation must be based on student growth on state assessments or a comparable measure of student achievement growth;
  - 20 percent of a teacher's evaluation must be based on locally-selected measures of student achievement that are determined to be rigorous and comparable across classrooms (decreases to 15% upon implementation of a statewide value-added growth model); and
  - 60 percent of a teacher's evaluation must be based on other measures of teacher/principal effectiveness (observations, survey tools, etc.).
- **D.C.**'s teacher evaluation system uses the following formula (for teachers for whom standardized assessment data are available):
  - A teacher's individual value-added student achievement data makes up 50 percent of the evaluation score.
  - A teacher's instructional expertise as measured by the Teaching and Learning Framework makes up 35 percent of the score.
  - A teacher's support for and collaboration with the school community make up 10 percent of the evaluation score.
  - The impact a teacher's school has on student learning over the course of the school year is 5 percent.
  - A final measure of professional requirements is included, although it is not weighted in the score. Points are deducted from the overall IMPACT score for teachers whose performance is below standards.

Figure 5. Measures and Methods in Teacher Performance Evaluations

STATE	Required components or standards	Weight of student outcomes (if specified)	Specific types of student achievement data required	Number of annual classroom observations (for non-probationary teachers)	Feedback conference required	Other measures or considerations
Arizona	All district evaluations must include three components: 1) classroom-level data, 2) school-level data and 3) teaching performance.	33-50%: Classroom-level data must account for at least 33% and school level data may account for up to 17%, for a total of no more than 50%	Classroom-level data may include state, district or school assessments, benchmark assessments and other standardized assessments, not teacher-made quizzes or tests. If not available, school, team or grade level achievement data may be used.	Multiple observations	No	For teachers with classroom level data available, additional school level measures such as aggregate state assessment data, AP/IB results, survey data and other school-level reliable measures are optional. For teachers without reliable classroom data, school-level measures are required.
Colorado	All district evaluations must measure five professional practices: 1) know content, 2) establish environment, 3) facilitate learning, 4) reflect on practice and 5) demonstrate leadership and student growth.	50%	1) measure of individually-attributed growth; 2) a measure of collectively-attributed growth whether on a school-wide basis or across grades or subjects; 3) when available, statewide summative assessment results; and 4) Colorado Growth model for subjects with annual state-wide summative assessment results available in two consecutive grades.	At least 1	No	Option to include reviews of teacher lesson plans or student work samples, surveys and other evidence.
Delaware	The state model measures 1) planning and preparation, 2) classroom environment, 3) instruction, 4) professional responsibilities and 5) student improvement.	Teacher cannot be rated "effective" overall if the student growth expectations for the teacher's students are not met. Schoolwide state test data are 30% of student improvement component; state test student cohort measures account for 20%. Teacher specific assessment measures are 50% of student improvement component.	For tested grades and subjects, scores on the statewide assessment and other measures of student learning that are rigorous and comparable across classrooms. For nontested grades and subjects, alternative measures of student learning such as scores on pre-tests and end-of-course tests, performance on English language proficiency assessments, and other measures that are rigorous and comparable across classrooms.	If "highly effective," 1, all others 2	Yes	Districts can develop and implement other assessment tools that measure student improvement, including assessments in other content areas, with state approval. Districts also may implement evaluations in addition to the state model. Among other options for evidence are: classroom management documents; examples of student work/assignments; and communication logs with parents.
D.C.	The District's system measures 1) student achievement (both individual and school value added), 2) instructional expertise, 3) commitment to school community and 4) professionalism.	50% individual value added measures and 5% schoolwide value added measures	Individual and school level value-added data for teachers in grades 4-8. Teacher-assessed data on tests other than statewide assessments and school level value-added data for other teachers.	5	Yes	No optional measures specified.
Florida	Districts design model with student growth requirements plus four domains: 1) classroom strategies and behaviors, 2) planning and preparing, 3) reflections on teaching and 4) collegiality and professionalism.	50% unless there are fewer than 3 years of data for a teacher, then 40%	State-adopted student growth measures for grades and subjects with state assessment data. District assessments for subjects and grades not covered by state test, and principal-set targets where no district tests available.	At least 1	No	System must include at least one "additional metric" of student performance to have a multi-metric evaluation for teachers in the year before a "milestone event." Parents must have an opportunity for input on teacher performance ratings.
Idaho	In addition to student growth, the state requires four domains: 1) planning and preparation, 2) learning environment, 3) instruction and use of assessment and 4) professional responsibilities.	At least 50%	Objective measures of student growth, as determined by the board of trustees.	Required but frequency not specified	No	Must include input from parents as a factor.
Illinois	TBD by September 2012	Must be a "significant factor"	TBD (by September 2012)	Required but frequency not specified	No	System must also consider the following: the teacher's attendance, planning and instructional methods; classroom management, where relevant; and competency in the subject matter taught.

STATE	Required components or standards	Weight of student outcomes (if specified)	Specific types of student achievement data required	Number of annual classroom observations (for non-probationary teachers)	Feedback conference required	Other measures or considerations
Indiana	In addition to a core professionalism rubric, the state's model rubric includes three domains: 1) purposeful planning and 2) effective instruction and 3) teacher leadership.	Must "significantly inform" the evaluation	Student assessment results from 1) statewide assessments; 2) methods for assessing growth for teachers in areas not measured statewide, including results from locally developed assessments and other tests.	Required but frequency not specified	No	Other performance indicators and model rubrics TBD by January 2012.
Louisiana	Half of every evaluation rating based on GO-Index (Growth Outcomes Index). The remaining half, the SITE-Index, is based on traditional evaluation techniques, such as classroom observations.	50%	Results from value-added assessment model determined by state board for grades where data are available. For grades, subjects and personnel for which value-added data are not available, the state board shall establish growth measures.	1	Yes	Additional local board criteria may be established based on job descriptions.
Maryland	Evaluation standards focused on student growth and professional practice. State framework on professional practice requires evaluation of 1) planning and preparation, 2) instruction, 3) classroom environment and 4) professional responsibilities.	Multiple measures of achievement to total 50%, no single measure counts for more than 35%	A single assessment may not be used solely as evidence of student growth. Districts choose measures based on a menu of approved options. If a statewide assessment is available for a teacher, it must be one of the measures used.	At least 2	No	Districts may include additional local priorities and measures such as scholarship, management skills, professional ethics and interpersonal relationships, with state approval.
Michigan	Student growth is specified, other standards TBD.	2013-14: 25% 2014-15: 40% 2015-16 at least 50%	TBD-Awaiting state Council on Educator Effectiveness recommendations. Student growth to be measured by national, state or local assessments and other objective criteria.	Multiple observations	No	System may include the following: instructional leadership abilities, teacher and pupil attendance, professional contributions, training, progress report achievement, school improvement plan progress, peer input, and pupil and parent feedback.
Minnesota	Only value added or student growth required by state. Local school board and teacher representatives agree on further detail.	35%	School boards must 1) use an agreed upon teacher value-added assessment model for grades and subjects where data are available and 2) establish state or local student growth measures where value-added data are not available.	Not specified	No	System must include option for teachers to present a portfolio demonstrating professional growth and teachers' own performance assessments based on student work samples. System must use longitudinal data on student engagement and connection and other student outcome measures aligned with curriculum.
Nevada	In addition to student achievement requirement, evaluations must include the following: classroom management skills, a review of lesson plans or grade book, whether the curriculum taught is aligned with standards, and whether the teacher is appropriately addressing the needs of students.	At least 50%	Information on pupil achievement maintained by the automated system of accountability information for the state must account for at least 50 percent of teacher evaluations.	Not less than 60 minutes per evaluation period	Yes	None specified.
New York	Evaluation system includes statewide student growth measures; locally selected measures of student achievement; teacher observations, school visits and other measures to provide teachers with detailed, structured feedback on professional practice.	40%: 25% state-developed measures; 15% locally selected measures by 2013-14	State Board of Regents to develop a value-added model; until then a mix of state assessments and locally-selected measures. For teachers without state assessment data, district-wide growth goal-setting with other standardized assessments from a state- approved list of options. Same measure may be used for both state assessments and locally selected measures subcomponents.	Required but frequency not specified	No	Optional additional element includes structured reviews of student work, portfolios, feedback from students, parents or others using structured surveys and teacher self-reflection. Menu of state-approved rubrics for assessing New York State Teaching Standards TBD.
Ohio	In addition to student academic growth requirements, measures of 1) professional goal-setting, 2) formative assessment of teacher performance and communication and 3) professionalism.	50%	State requires, where applicable, value added progress measures, which includes performance on statewide achievement tests.	At least 2	No	Evaluation includes a self-assessment for all components. Experienced teachers who perform at an "accomplished" level may choose to complete a professional project as part of the annual evaluation.

STATE	Required components or standards	Weight of student outcomes (if specified)	Specific types of student achievement data required	Number of annual classroom observations (for non-probationary teachers)	Feedback conference required	Other measures or considerations
Oklahoma	In addition to student achievement, the state is considering three framework options for teacher evaluation: Danielson's Framework for Teaching; Marzano's Causal Teacher Evaluation Model, and Tulsa's TLE Observation and Evaluation System.	35% growth measures and 15% other academic measures	Measures of academic growth using multiple years of standardized test data. Where there is no state-mandated testing measure, objective measures including student performance on unit or end-of-year tests and overall school growth.	Required but frequency not specified	No	None specified.
Rhode Island	State model includes 1) student learning, 2) professional practice and 3) professional responsibilities.	Student learning is predominant component through matrix scoring model.	Rhode Island growth model to assess teachers of reading and math for grades 3-7. All teachers will set specific, measurable student learning objectives.	At least 4	Yes	Teachers engage in self-assessment and prepare professional growth plans in addition to participating in setting student learning objectives (SLOs).
Tennessee	In addition to student growth, state uses the TAP rubric, an observation tool that includes 26 indicators of teaching skills focused on 1) planning, 2) environment, 3) professionalism and 4) instruction.	35% student achievement growth; 15% other achievement measures	Evaluations based on student growth data from the Tennessee Value-Added Assessment System (TVAAS) or other comparable measure of growth. Teachers with TVAAS who teach grades 4-8 may choose among the following achievement measures: state assessments, schoolwide TVAAS, ACT/SAT suite of assessments, "off the shelf" assessments and completion/ success in advanced coursework. In addition to the measures already listed, secondary teachers with TVAAS may also choose the following: AP/IB/NIC suites of assessments, graduation rates, postsecondary matriculation and 9th grade promotion to 10th grade.	4	Yes	Mandatory criteria include review of prior evaluations and classroom observations.

These formulaic weighting of sets of criteria are not the only approach to defining the measures to be included in performance-based evaluations. **Delaware**'s DPAS II system formula is organized around five key evaluation components or topics: 1) planning and preparation, 2) classroom environment, 3) instruction, 4) professional responsibilities and 5) student improvement. Teachers are rated highly effective, effective, needs improvement or ineffective in each individual area. A summative rating depends on the number of effective or not effective ratings teachers receive in each of the individual component areas. All other components aside, if a teacher does not meet or exceed student growth requirements in Delaware, the teacher cannot be rated any higher than needs improvement overall, regardless of ratings in the other four components.

Assigning Teacher Effectiveness Ratings

Underlying the basic operating formulas we've just described is a more complicated, but critically important, set of rules for how the various components of a teacher's annual evaluation are arrived at, combined and defined into categories of effectiveness.

Beyond identifying the numbers and names of performance levels, most states have not decided much of the remaining specifications. The devilish details of transferring the concept of performance-based teacher evaluation and the outline of components and multiple measures of performance into meaningful rating categories have become the business of state technical advisory committees, stakeholders and evaluation and measurement experts.

Figure 6. Teacher Performance Levels/Categories

STATE	Number of teacher evaluation performance levels	Category names
Arizona	Not specified	N/A
Colorado	3	Highly effective, effective and ineffective[22]
Delaware	4	Highly effective, effective, needs improvement and ineffective
D.C.	4	Highly effective, effective, minimally effective and ineffective
Florida	4	Highly effective, effective, needs improvement and unsatisfactory
Idaho	2	Minimum of two categories to address proficient and unsatisfactory
Illinois	4	Excellent, proficient, needs improvement and unsatisfactory
Indiana	4	Indiana Teacher Effectiveness Pilot calls for highly effective, effective, improvement necessary and ineffective
Louisiana	2	Effective and ineffective[23]
Maryland	3	Maryland pilot program using highly effective, effective and ineffective
Michigan	Not specified	State requires that "multiple rating categories" are used
Minnesota	Not specified	N/A
Nevada	4	Highly effective, effective, minimally effective and ineffective
New York	4	Highly effective, effective, developing and ineffective
Ohio	4	Accomplished, proficient, developing and ineffective
Oklahoma	5	Superior, highly effective, effective, needs improvement and ineffective
Rhode Island	4	Highly effective, effective, developing and ineffective
Tennessee	4	Advanced, proficient, developing and unsatisfactory

22 Colorado's legislation specifies three rating categories. Draft regulations indicate there will be four.

23 Louisiana's legislation only identifies two classifications. Draft materials suggest there will be additional categories.

Taken together, there still seem to be some decisions of consequence to be made around simply choosing the number of evaluation classifications. The only clear right answer at present on the number of performance levels seems to be more than two. We think four performance levels may be the better option for rating teachers using subjective observation protocols, and five may be optimal for overall ratings. On the observation side, an even number like four forces raters to differentiate, avoiding the tendency to lump everyone into the safe middle provided by an odd number. On the summative side, where there presumably will be a set of rules for combining the various measures rather than rater discretion, five categories provides more room for real differentiation, particularly in distinguishing the truly ineffective and the most outstanding performers. However, there is no definitive right answer.

How do states translate their evaluation measures into definitions of teacher performance? Most states don't have this worked out yet. But for some of the states where new evaluation systems are up and running, such as **D.C.** and **Delaware**, we can take a look at developing approaches.

Figure 7. Defining Teacher Effectiveness: State Examples

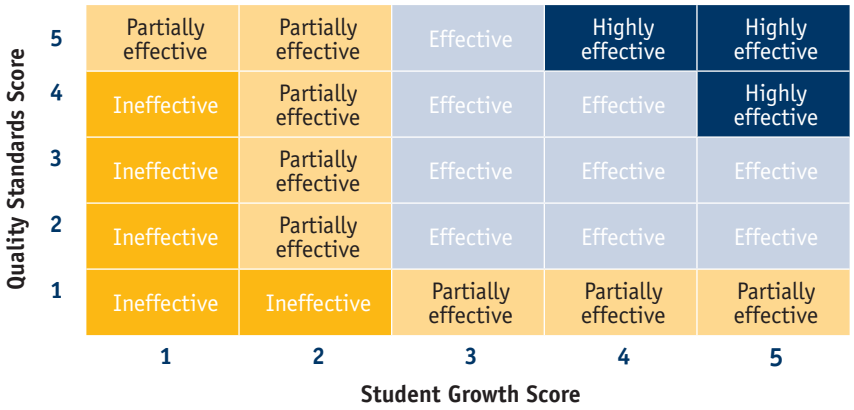
STATE	Method of scoring evaluation	Definitions of teacher performance levels
Delaware	Delaware assesses five evaluation components. On component 1-4 (non-student growth components) teachers receive a rating of distinguished, proficient, basic or unsatisfactory. For the student achievement component 5, assessment will yield a combined 100-point scale: 80-100 points will be awarded a rating of exceeds. Scores of 50-79 points will be rated as satisfactory on this component, and a score below 50 will be rated as unsatisfactory.	**Highly Effective** means a teacher has earned a satisfactory component rating in four of the five components including a rating of exceeds for the student growth component. **Effective** means a teacher has earned a satisfactory rating in at least three of the five components, including student growth, but the teacher does not meet the student growth requirements for a highly effective rating. **Needs Improvement** means a teacher has earned one or two satisfactory ratings out of the five components including a satisfactory rating in the student growth component or a teacher has earned three or four satisfactory ratings out of the five components and the teacher has earned an unsatisfactory rating in student growth. **Ineffective** means a teacher has received zero, one, or two satisfactory ratings out of the five components, and the teacher has received an unsatisfactory rating in student growth or a teacher's overall summative evaluation rating is needs improvement for three consecutive years. In this case, the teacher's rating is re-categorized as ineffective.

STATE	Method of scoring evaluation	Definitions of teacher performance levels
DCPS	DCPS generates an IMPACT score for each teacher, which ranges from 100-400. Each component (except Core Professionalism) of the assessment is measured on a scale of 1-4 (with 1 being the lowest score and 4 being the highest score). Each component score is multiplied by the percentage weight of that component in the evaluation system, which creates weighted scores that are added together to arrive at a total score between 100 and 400. That score is adjusted by 10 to 20 points (downward) if professionalism standards are not met.	**Highly Effective (350-400)**: This rating signifies outstanding performance. **Effective (250-350)**: This rating signifies solid performance. **Minimally Effective (175-250)**: This rating signifies performance that is below expectations. Individuals who receive a rating of minimally effective for two consecutive years will be subject to separation from the school system. **Ineffective (100-175)**: This rating signifies unacceptable performance. Individuals who receive this rating will be subject to separation from the school system.

Unlike Delaware and D.C., which have come up with composite scale scores that translate into a specific teacher performance classification, **Rhode Island**'s formula uses a matrix approach to combine performance on multiple evaluation components. This method uses a series of look-up tables to identify a rating for teachers on professional practice and professional responsibilities and then combine those scores. Another matrix helps identify ratings related to student learning objectives and student growth requirements (where available). The state then uses another matrix that combines professional standards and student learning standards to identify an overall rating for a teacher.

Figure 8. Example of a Matrix for Classifying Teacher Effectiveness

These kinds of look up tables are user-friendly and can add a level of transparency to the way teacher ratings from various instruments, observations and other criteria included in evaluations are scored.



State Use of Evaluation Results

Consistently across states, the articulated goals of new teacher evaluations are to improve instruction, to better design professional development to meet teachers' needs and to increase the effectiveness of the teacher workforce. However, to meet these ambitious goals, teacher evaluations cannot stand alone. States need to put in place supporting policies to ensure that high-quality teacher evaluation data are used to make important decisions that will matter for teacher effectiveness, including what to do about low-performing teachers.

Several components of teacher policy ought to be in place if performance-based teacher evaluations are going to help districts better handle the problem of low performance:

1. States should have policies for assisting teachers who receive poor evaluations. However, these assistance policies should be time limited.
2. States should be explicit that ineffective classroom performance is grounds for teacher dismissal.
3. Teachers should not be able to appeal their performance rating. Certainly a teacher who is terminated for poor performance should have an opportunity to appeal. But, unlike prevailing practice today, the state should ensure that this appeal occurs within a reasonable time frame and is distinct from the due process procedures for teachers dismissed for felony or morality violations or dereliction of duties. Further, the state should ensure decision-making remains in the hands of those with educational expertise.

Figure 9. Consequences for Poor Teacher Evaluation Ratings

With very few exceptions, the states rethinking and revamping their teacher evaluation systems are specifying in legislation and regulations a set of expectations and responsibilities that school systems must set in motion when teachers receive poor evaluations. The most promising policies on this front spell out both the kinds of interventions required and a specific time period within which ineffective teachers should have an opportunity to demonstrate improvement or be dismissed; policies we find noteworthy in this regard are highlighted below.

STATE	State policy for assisting teachers who receive poor evaluations	Teachers are eligible for dismissal based on poor evaluations	State policy for dismissing ineffective teachers
Arizona	A low evaluation must include recommendations as to areas of improvement. Assistance and opportunities must be provided for the certificated teacher to improve performance. After a reasonable period of time, the designee must follow up with the teacher to ascertain whether that teacher is demonstrating adequate classroom performance.	No	
Colorado	Each teacher must be provided with an opportunity to improve effectiveness through a teacher development plan. School districts must ensure that a teacher who objects to a rating has an opportunity to appeal, in accordance with a fair and transparent process developed, where applicable, through collective bargaining. For non-probationary teachers, a remediation plan must be developed by the district and must include professional development opportunities. The teacher must be given a "reasonable period" to improve.	Yes	Colorado specifically identifies classroom ineffectiveness as grounds for dismissal. For teachers who receive a performance rating of ineffective, the evaluator shall either make additional recommendations for improvement or may recommend dismissal.

STATE	State policy for assisting teachers who receive poor evaluations	Teachers are eligible for dismissal based on poor evaluations	State policy for dismissing ineffective teachers
Delaware	Teachers who receive an overall rating of needs improvement or ineffective on the summative evaluation, or a rating of unsatisfactory on any appraisal component regardless of the overall rating, must be put on an improvement plan.	Yes	Teachers with two consecutive years of ineffective ratings or who earn a combination of ineffective and unsatisfactory ratings for three consecutive years are considered to have a pattern of ineffective teaching and are eligible for dismissal.
D.C.	Those who are rated minimally effective are encouraged to take advantage of professional development opportunities provided by DCPS.	Yes	DCPS ensures that teacher ineffectiveness is grounds for dismissal. Individuals who receive ineffective ratings are "subject to separation from the school system."
Florida	If a teacher receives an unsatisfactory evaluation, the evaluator must make recommendations as to specific areas of unsatisfactory performance and provide assistance in helping to correct deficiencies within a prescribed period of time.	Yes	Florida ensures that teacher ineffectiveness is grounds for dismissal. All new teachers are placed on annual contracts and the state requires that such contracts are not renewed if a teacher's performance is unsatisfactory. An annual contract may not be awarded if the teacher has received two consecutive annual performance evaluation ratings of unsatisfactory, OR two annual performance ratings of unsatisfactory within a three-year period, OR three consecutive annual performance evaluation ratings of needs improvement or a combination of needs improvement and unsatisfactory.
Idaho	Not specified. Districts can propose procedures to provide remediation in those instances where remediation is determined to be an appropriate course of action.	No	
Illinois	Those who receive a rating of needs improvement must be placed on professional development plan to address those areas. Those rated unsatisfactory must be placed on a remediation plan.	Yes	Illinois specifically identifies classroom ineffectiveness as grounds for dismissal. For teachers placed on remediation plans for poor performance that receive a subsequent unsatisfactory performance rating within three years, the school district may forego remediation and seek dismissal.
Indiana	Not specified	Yes	Indiana ensures that teacher ineffectiveness is grounds for dismissal. A tenured teacher reverts to probationary status if the teacher has received a rating of ineffective on an evaluation and can be subject to contract cancellation for a rating of ineffective in the year immediately following the teacher's initial rating of ineffective.
Louisiana	Any teacher not deemed effective will be placed in an intensive assistance program and then must be formally re-evaluated. Program must include an expected time line for achieving objectives; must not exceed two years.	Yes	If at the end of intensive assitance program, a teacher does not complete the program or is still deemed ineffective based on evaluation, the school district is allowed to initiate termination proceedings.
Maryland	An unsatisfactory evaluation must include at least one observation by someone other than the immediate supervisor. Teachers may appeal overall ratings of unsatisfactory; the burden of proof is on teacher.	No	
Michigan	Teacher must be given “ample opportunities for improvement.”	Yes	Michigan identifies classroom ineffectiveness as grounds for dismissal. If a teacher is rated as ineffective on 3 consecutive annual year-end evaluations, the district shall dismiss the teacher.
Minnesota	Districts must give teachers not meeting professional teaching standards support to improve through a teacher improvement process that includes established goals and timelines.	No	
Nevada	Districts are responsible to ensure that a “reasonable effort” is made to correct deficiencies upon the request of a teacher who needs assistance.	Yes	Nevada ensures that teacher ineffectiveness is grounds for dismissal. All post-probationary teachers will return to probationary status if they receive two consecutive years of unsatisfactory evaluations.

STATE	State policy for assisting teachers who receive poor evaluations	Teachers are eligible for dismissal based on poor evaluations	State policy for dismissing ineffective teachers
New York	If a teacher is rated developing or ineffective, the school district is required to develop and implement a teacher or principal improvement plan.	Yes	Tenured teachers with a pattern of ineffective teaching or performance, defined as two consecutive annual ineffective ratings, may be charged with incompetence and considered for termination through an expedited hearing process.
Ohio	Each teacher must be provided with a written report of the results of the teacher's evaluation that includes specific recommendations for any improvements needed in the teacher's performance, suggestions for professional development that will enhance future performance in areas that do not meet expected performance levels, and information on how to obtain assistance in making needed improvements.	Unclear	Unclear[24]
Oklahoma	All teachers who receive ratings of needs improvement or ineffective must be placed on comprehensive remediation plans and provided with instructional coaching.	Yes	Oklahoma ensures that teacher ineffectiveness is grounds for dismissal. Teachers rated as ineffective for two consecutive years; needs improvement for three years; or for those who do not average at least an effective rating over a five-year period shall be dismissed.
Rhode Island	Teachers who are rated as developing or ineffective at the end of the year will be placed on an individual development plan and will work with an improvement team to assist them with their development over the course of the following year. The teacher's district will identify personnel actions that may occur if he or she does not adequately improve his or her performance.	Yes	Rhode Island explicitly makes teacher ineffectiveness grounds for dismissal by requiring districts to dismiss all educators who are rated ineffective for two consecutive years.
Tennessee	Not specified.	Yes	Tennessee explicitly makes teacher ineffectiveness grounds for dismissal. Tennessee specifies that tenured teachers who receive two consecutive years of below expectations or significantly below expectations performance ratings are returned to probationary status, making them eligible for dismissal.

Beyond specifying improvement and dismissal policies, there are a variety of ways that states are making linkages between teacher evaluation results and teacher effectiveness strategies. These include eliminating tenure and tying effectiveness to tenure, revising "last in, first out" policies, and providing principals with more discretion to hire staff. These broader teacher effectiveness strategies also include tying professional development and compensation to evaluation results.

- **Eliminating tenure policies. Florida**, for example, has discontinued its policy of tenure and bases annual teacher contracts on performance in the classroom. To be awarded or renew an annual contract, a teacher must not have received any of the following evaluation ratings: two consecutive annual performance evaluation ratings of unsatisfactory; two annual performance evaluation ratings of unsatisfactory within a three-year period; or three consecutive annual performance evaluation ratings of needs improvement or a combination of needs improvement and unsatisfactory.
- **Awarding tenure based on effectiveness**. Probationary teachers in **Colorado** must earn three consecutive effective ratings to become nonprobationary. Veteran, or nonprobationary, teachers who receive two consecutive ineffective ratings return to probationary status and have a year to improve or face termination. In **Delaware**, the state now requires that probationary teachers must show two years of satisfactory student growth—evidenced by satisfactory ratings in the "student improvement" component of the teacher appraisal process—within a three-year period before they earn tenure.

24 Due to the pending referendum on its education reform bill, Ohio currently has two versions of its education code pertaining to teacher evaluation on its books. One version addresses removal of poorly performing teachers, but the other does not.

Michigan's recently passed tenure legislation has increased the probationary period for new teachers to five years. The state also now articulates that a teacher has not successfully completed this probationary period unless he or she has been rated as effective or highly effective on the three most recent annual performance evaluations.

Tennessee has recently increased its probationary period to five years and now requires probationary teachers to receive an overall performance effectiveness rating of above expectations or significantly above expectations during the last two years of the probationary period. A tenured teacher who receives two consecutive overall ratings of below expectations or significantly below expectations may be reverted to probationary status until they receive two consecutive ratings of above expectations or significantly above expectations.

- **Tying layoff policies to teacher effectiveness, not just seniority. Indiana** has ended the state's "last in, first out" policy for reductions in force, which resulted in laying off the newest teachers first, regardless of effectiveness. Now, a teacher's performance must be the first criterion used; seniority may be considered among other criteria for teachers in the same performance category. **Colorado**, **Florida** and **Michigan** have also ended 'last in, first out" practices.
- **Providing principals with discretion to hire. Colorado**'s law prohibits a teacher from being assigned to a school without the principal's consent. These are important strides in giving school leaders the authority and at least some flexibility to make choices related to building an effective teaching staff. The latest contract between **DCPS** and the Washington Teachers Union represents significant advancement on this issue. One of the key components of the contract, which went into effect July 2010, is mutual consent regarding teacher hiring, meaning that the teacher and the school must agree for a teacher to get the job. It applies regardless of tenure, so principals may now staff their schools based on the most qualified candidates.
- **Tying teacher compensation to effectiveness.** Starting in 2014, **Florida** will require that districts tie teacher compensation to teacher performance. A teacher determined to be highly effective will receive a salary increase that must be greater than the highest annual salary adjustment available to that individual through any other salary schedule adopted by the school district. A teacher determined to be effective will receive a salary increase between 50 and 75 percent of the annual salary increase provided to a highly effective employee.

  Indiana's new law limits the extent to which teacher salary can be based on seniority and education level and specifies that teachers with ineffective ratings are not eligible for pay raises. **Tennessee** requires local districts to develop differentiated pay plans; those plans may include pay based on performance. If a district chooses to include a performance component, it must be "based on gains in student academic achievement" and "be criterion-based so that everyone meeting a previously agreed-upon standard earns that award." The amount of the award for effective teaching is decided at the local level, but the state requires that the amount be in the thousands, not hundreds of dollars – incentives significant enough to matter to teachers.
- **Streamlining dismissal policies. Indiana** has adopted a new streamlined appeals process. Appeals are made to the local school board, whose decision must be reached within 30 days and is final. In **New York**, tenured teachers and principals with a pattern of ineffective teaching or performance, defined as two consecutive annual ineffective ratings, may be charged with incompetence and considered for termination through an expedited hearing process.
- **Providing timely feedback and aligning professional development with effectiveness ratings.** Developing improvement plans for teachers who receive poor evaluations, providing timely feedback

and designing professional development to be aligned with evaluation results are key functions of performance-based teacher evaluation systems. This requires more than just giving teachers a copy of their evaluation forms. **Michigan** requires that annual performance evaluations provide teachers with "timely and constructive feedback." In addition, the state requires that evaluations be used to inform relevant coaching, instructional support, and professional development.

Delaware requires that teachers receive feedback from their evaluations during an end-of-year conference. The state also specifies that findings shared during the conference should be used to inform a teacher's future professional development activities. For teachers on improvement plans, required professional development activities must be aligned with findings from teachers' evaluations.

Rhode Island requires that all teachers receive written, detailed feedback that informs recommendations for professional growth. The state also specifies that evaluation systems be designed to provide "agreement between the evaluation analysis and the identified goals and improvement expectations that inform professional development."

- **Tying certification to effectiveness.** There are two states where a teacher's continued licensure is linked to performance evaluations. In **Louisiana**, Act 54, passed in May 2010, specifies that teachers must meet a standard for effectiveness, established by the state, based on a performance evaluation that includes growth in student achievement using value-added data. Teachers must meet the standard for effectiveness for three years during their initial certification or renewal period to be issued a certificate or have their certificate renewed. In **Rhode Island**, any teacher with five years of ineffective ratings would not be eligible to have his or her certification renewed by the state.

Figure 10. Teacher Effectiveness Policies Tied to Evaluation Results

STATE	Tenure	Dismissal	Certification	Compensation connected to teacher evaluation results	Effectiveness factored into layoff decisions
Arizona				✓	
Colorado	✓	✓			✓
Delaware	✓				
D.C.		✓		✓	✓
Florida	✓[25]	✓		✓	✓
Idaho					
Illinois	✓[26]	✓			✓
Indiana	✓	✓		✓	✓
Louisiana			✓	✓	
Maryland					
Michigan	✓	✓		✓	✓
Minnesota				✓[27]	
Nevada	✓	✓		✓	✓[28]
New York	✓	✓			
Ohio	✓	✓		✓	
Oklahoma		✓			✓
Rhode Island	✓	✓	✓		✓
Tennessee	✓	✓		✓	

25 Florida has effectively eliminated tenure; annual contracts based on performance.

26 Policy applies only to districts with fewer than 500,000 students, effectively making this policy not apply to Chicago Public Schools.

27 Q Comp is based on various measures of student achievement, but it is not clear whether the program will incorporate the new growth measures required by the state.

28 In Nevada, new legislation ensures that seniority will not be the sole factor in determining which teachers are laid off during a reduction in force. Among other things performance evaluations "may" be considered.

Early Lessons

on the Road to Performance-Based Teacher Evaluation

It is still early to critique the state of the states on teacher evaluation. Most policies are still very new, many of the details have yet to be determined, and few systems are up and running. But the changing landscape of teacher evaluation policy provides an opportunity to reflect on some of the early lessons. For those states already down the road and those starting down the path, we can report on how states are doing, what some of the pitfalls are and what strategies seem worth emulating. The thinking below is drawn both from our comparative analysis of state policies as well as our own experience advising states and districts on these issues.

1. Teacher effectiveness measures don't have to be perfect to be useful.

Are emerging teacher effectiveness measures perfect? No. But they are a marked improvement on evaluation systems that find 99 percent of teachers effective with little attention to a teacher's impact on students and offer little meaningful information on teachers' strengths, weaknesses and professional development needs. Do the new systems coming online have the potential to shed light on effective practice and improve teaching and learning? Yes.[29]

Student growth and value-added methodologies are still emerging. However, examining student achievement as a metric for assessing teacher effectiveness, even if measurement is imperfect, represents a big step forward.

Indeed, we set the whole enterprise up for failure if we attach unrealistic expectations to the exact precision of every measure – and doom ourselves to the alternative of doing very little to measure and examine teacher performance. At the same time, the reality of measurement and limits to teacher control over student outcomes do argue for measured caution in developing teacher effectiveness policies.

29 For a good discussion of why teacher performance measures need to be good, but not perfect, and how measures of teacher effectiveness stack up against other measures of "success" see "Passing Muster: Evaluating Teacher Evaluation System" Brown Center on Education Policy at Brookings Institution, April 2011, at: http://www.brookings.edu/reports/2011/0426_evaluating_teachers.aspx. For example, most test-based measures of teacher effectiveness range from .2 to .6. By comparison, SAT correlations with measures of college success are .35 (for SAT combined verbal and math as predictor of first year GPA). Those correlations may be considered modest but almost all selective colleges weight SAT scores (and many very heavily) in admission decisions.

2. Insistence on comparability of measures for all teachers could cripple evaluation efforts.

As states are looking at implementing performance-based teacher evaluation systems, one of the challenging issues they face is developing measures of student achievement growth in grades and subjects for which *consistent* statewide assessment data are not available. How can states measure growth and attribute the value teachers add to student learning under these circumstances?

This is no doubt a complicated issue. But the pursuit of comparable data may be an issue where states are a bit too hung up. The drive to identify or develop comparable measures for teachers regardless of grade or subject taught is understandable, but the more important emphasis ought to be on *fair and valid* measures.

Measurement that varies by type of teacher – music versus biology, or social studies versus vocational education – is a kind of "inconsistency" that we can not only live with, but is appropriate in evaluating teachers. We need to abandon the lock step mentality that has controlled too many aspects of the teaching profession. Comparability of all measures isn't the ultimate goal; *fair, rigorous and appropriate measures* of teacher performance are the bottom line. Developing such measures for grades and subjects for which there are no statewide measures is a valuable process.

This isn't to argue that where states (and districts) have comparable data across sets of teachers, those data shouldn't be used to measure student growth and teacher effectiveness in a way that maintains comparability. In fact, states should insist on this. But where those measures don't exist, the choices aren't between developing statewide tests for every grade and subject and throwing out the whole project because we can't measure growth and evaluate teachers all in exactly the same way.

3. Designing measures of student growth for non-tested grades and subjects is an important challenge facing states.

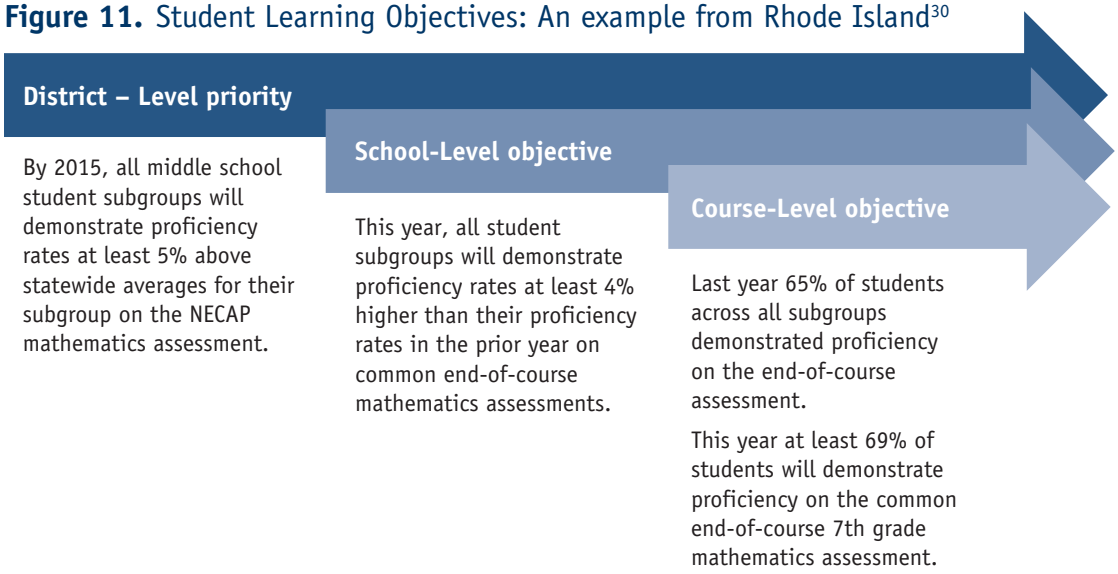
Thinking about the full complement of teachers – including K-2, social studies, special education and non-core subject area teachers – states face a challenge of how to develop fair and rigorous measures of student growth and achievement that can be used to evaluate the performance of teachers for whom state standardized achievement data do not exist.

States are approaching this challenge in a variety of ways. Some are contemplating new tests in subject areas such as social studies or grades (such as K-2) for which statewide testing is not currently available. **Delaware**, for example, is in the process of developing, with input from teachers across the state, lists of approved measures for teachers of every grade and subject. The state will utilize a pre/post assessment model to determine growth, although these assessments may look very different for different subjects. Approved assessments will include commercially-available externally-developed measured, as well as validated teacher-made assessments. States are also developing lists of appropriately rigorous local district measures that can be used to develop growth measures. **Florida** is exploring the use of item banks that can be used to assess grades and subjects where statewide assessment data are not available. **Tennessee** is using school-level value-added measures where individual classroom measures don't exist.

Another strategy states are exploring as a potential solution to the lack of value-added or growth data for many teachers, is the development of student learning objectives (SLOs).

SLOs are achievement goals set for groups of students based on a mix of available, objective state and local measures that are developed by teams of administrators, grade-level teams or groups of content-alike teachers. In **Rhode Island**, where SLOs are part of the state's evaluation model, these measures are developed in cooperation with teachers who are responsible for helping specific students meet specific learning objectives. (Typically a teacher will be responsible for 2-4 objectives.) All teachers develop SLOs, and in grades and subjects where no statewide growth data are available, these measures take on greater weight in teacher evaluations.

Figure 11. Student Learning Objectives: An example from Rhode Island[30]



Done well, SLOs can provide meaningful benchmarks to gauge teacher performance. Done poorly – with low standards or misaligned targets – they may differ little from current low-quality evaluation metrics.

4. States shouldn't lose sight of the importance of classroom observations.

While there is a great deal of attention focused on linking value-added and student growth results to teacher evaluation, it is equally important to gather evidence observing behavior – what teachers do and what students are learning in the classroom – during classroom observation. The criticism of many current evaluation systems is not just their failure to take student learning into account, but their failure to include high-quality classroom observations.

30 From the Rhode Island Model: Guide to Evaluating Building Administrators and Teachers 2011-12, p.46.

Assessing the Quality of Classroom Observation Instruments

1. **Do the evaluation criteria focus on behaviors that can be observed?** A strong observation rubric should focus almost exclusively on teacher practices and student behaviors that can be observed in the classroom. While other criteria are not without merit, they may call for too much subjectivity and guesswork on the part of the evaluator.
2. **Do the criteria incorporate teacher and student evidence, requiring evaluators to look for direct evidence of student engagement and learning as well as evidence of teacher performance?** Too many rubrics fail to take into account evidence of student behavior. It is not enough to ask for evidence of teacher performance without looking for evidence that the teacher's performance is having an appropriate impact on the students.
3. **Is the number of standards or elements an observer is expected to evaluate manageable?** A classroom observation can easily become an unmanageable and cursory checklist review if the evaluator has too many standards to assess in a relatively short classroom visit. The evidence an evaluator is expected to collect during classroom observations must be easily accumulated during a typical lesson or over multiple visits and should be focused on what matters most.

New research is encouraging on this front. Well-designed and executed classroom observations can be effective at identifying the effectiveness of teachers, particularly teachers at the top and bottom ends of the distribution. Recent research also finds that good evaluations impact teacher effectiveness[31] – that is, evaluations don't have to be just summative report cards. If done well, they can indeed be formative tools that drive teacher improvement.

There are two key factors to a strong observation system. First, instruments must measure the right behaviors. Unfortunately, many observation rubrics are filled with vague descriptions of teacher practices, which may or may not have any relationship to student outcomes. Second, evaluators must be well trained to utilize the instrument so that results will be valid.

States with statewide evaluation systems or state-developed models are adopting standard observation protocols. The **District of Columbia**'s IMPACT system uses standards protocols that are implemented by trained and certified evaluators and focus on concrete and observable classroom practices that correlate to student gains. Importantly, the system also focuses on providing teachers with feedback that is meaningful for improving practice.

It is clear that performance-based evaluations will require more from evaluators and observers of teacher performance than they have in the past. States will need to make huge investments in training, as even the best evaluation system will be crippled by poor implementation. The need for training represents an enormous undertaking for the states. It is no less daunting a task than training an army, given the range of personnel involved, including principals, assistant principals, department heads and teams of peer evaluators.

31 Kane, Taylor, Tyler and Wooten. 2011 "Evaluating Teacher Effectiveness: Can classroom observations identify practices that raise achievement?" *Education Next*, Vol 11, No.3

Among the states NCTQ examined, states vary in their expectations of evaluators. Some states, such as **Oklahoma** and **Tennessee** require that peer evaluators participate in training and certification. **D.C.**'s Master Educators are practitioners who serve as impartial teacher evaluators, and conduct classroom observations without any knowledge of the IMPACT scores that teachers receive from their administrators. Master educators also undergo training on the district's Teaching and Learning Framework.

Training needs to go beyond the observation itself to include follow-up activities. Whether principals or peers, reviewers will need training that prepares them to have meaningful discussions about what has been observed, particularly when these conversations involve significant weaknesses. Both **Colorado** and **Michigan** specify in legislation the state's plans to monitor and evaluate the evaluators for the purposes of examining the consistency of teacher ratings within the parameters of the state or district evaluation model.

Figure 12. State Policies Assigning Responsibility for Conducting Evaluations

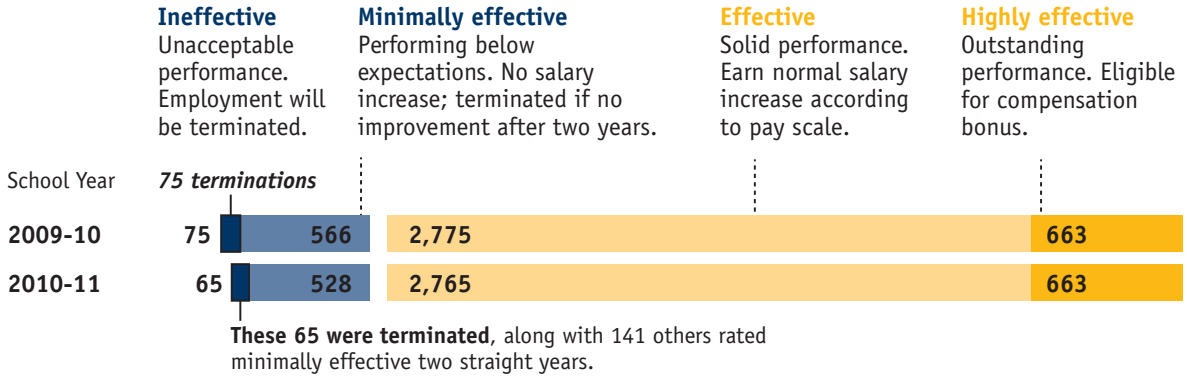
STATE	Who is responsible for conducting teacher observations/evaluations?
Arizona	Teachers designated by each school district's governing board
Colorado	Principals or administrators, who must be evaluated for compliance with the district's evaluation system
Delaware	Credentialed evaluators, who are trained and certificated and are usually the supervisor of the teacher
D.C.	Three of the observations are conducted by an administrator, and two are conducted by an impartial third-party "Master Educator"
Florida	The individual responsible for supervising the teacher
Idaho	Administrators, who must be trained
Illinois	Administrators, who must complete a training program and whose ratings have been determined by an independent observer to align to the requirements established by the State Board
Indiana	Individuals with demonstrated records of effective teaching and principal approval to conduct evaluations
Louisiana	School principals or vice principals, or respective supervisory-level designees
Maryland	Trained evaluator
Michigan	School administrator or designee
Minnesota	Individuals trained and qualified, such as school administrators
Nevada	Administrators
New York	The building principal, or his or her designee
Ohio	One or more of the following: a superintendent, assistant superintendent or principal; a vocational director or supervisor; and/or a person designated to conduct peer reviews through a collective bargaining agreement
Oklahoma	Principal, assistant principal or other trained certified individual designated by district board
Rhode Island	Principals, assistant principals department chairs or other instructional leaders
Tennessee	Principals, who must be trained and certified as evaluators

5. In addition to providing actionable feedback to all teachers, perhaps the most useful initial capacity of new evaluations will be to discern the most and least effective teachers.

The precision of growth and value-added data may not be at a very high level of sophistication, but that doesn't mean these data should be discounted. According to some recent research findings, replacing even the lowest performing 8 percent of teachers with an *average* teacher would put the U.S. on par with top performers on international math and science tests.[32]

Take the second-year data from **D.C.**'s IMPACT system. It resulted in the top 663 teachers (16 percent of the workforce) being eligible for bonuses of up to $25,000. That approximates pretty closely to what economists estimate is the percentage of truly exemplary teachers on the average district payroll.

FIGURE 13: Results from DC's IMPACT SYSTEM



About a third of these teachers earned this distinction for the second year in a row and is now eligible for a permanent base pay increase of up to $20,000. On the flip side, 206 teachers are being dismissed for poor performance (5 percent of the workforce), up from 135 last year. Of the 206 teachers let go this year, 65 were rated as ineffective, while another 141 were rated minimally effective for the second consecutive year. Just over half of those teachers rated minimally effective a year ago were rated effective this year, and 3 percent were rated highly effective.

What may be most the most interesting indicator of how good a job IMPACT is doing at identifying teachers who need to go is the response to the new policy in 2011 allowing principals to make exceptions to dismissing teachers who received a minimally effective rating two years in a row. It could have been a major loophole. But, in the end, the jobs of only four employees (and just two teachers) were saved with the change. Principals largely agreed that the teachers IMPACT said should be let go really needed to go.

32 Hanushek, Eric A. 2010. "The Economic Value of Higher Teacher Quality." National Bureau of Economic Research, Working Paper 16606, December.

6. Stakeholder input is important – but bold leadership is more important.

Nothing about building a truly effective teaching force is going to come easy. The reality is that teacher reform is being met with unparalleled, vocal opposition. In anticipation of such opposition, states need to look beyond current constituencies to achieve the necessary momentum. Many states are using advisory committees to flesh out the details of teacher evaluation not articulated by statute. While it is critically important to have stakeholder voices represented, it must be balanced with the need for technical expertise in building these systems.

States like Massachusetts and Colorado have had mixed results, with the recommendations from their advisory committees seeming to back-pedal from more ambitious and rigorous expectations. In Massachusetts's case, there were 45 people on the advisory committee, a large number indeed to reach consensus on any sort of reform. Other states, including **Rhode Island**, **Louisiana** and **Delaware**, have used technical advisory committees to provide expert input while using other mechanisms—including town hall meetings, on-line surveys and other committees—to receive stakeholder input.

There is also a need for real leadership to support not just the passage of legislation but implementation as well. Senator Michael Johnston has been a tireless advocate for **Colorado**'s efforts. Many states have reorganized their Education Departments, creating offices and positions specifically dedicated to educator effectiveness. In **Louisiana**, for example, teacher evaluation is housed in the Human Capital Office.

Finally, more so than any other strategy described herein, success may be dependent on an effective and proactive communication plan. Teachers are understandably worried about changing evaluation policy, and it is important that they be well-informed and have access to good sources for reliable information. Parents too need to understand the changes that are occurring. **Tennessee** SCORE, an education reform advocacy group, has played a central role in coordinating communications related to all aspects of that state's Race to the Top plan.

7. State review and approval of district evaluations may not be an adequate approach to ensuring quality and rigor.

State approval sounds like a good idea in states that leave it to districts to design a performance-based teacher evaluation system. But how realistic is it?

Florida is one of only three states that requires state review and approval of district evaluation plans. But it arguably had that same oversight before its new 2011 legislation and districts routinely ignored the state requirements for teacher evaluation.

The Q Comp system in **Minnesota** was a trailblazer for linking teacher evaluation and teacher compensation. But the details for Q Comp were left to the districts and research on the program has shown that almost all participating teachers were rated effective and getting bonuses in participating districts. The system didn't differentiate teacher performance. There is good reason to be skeptical that state review of all district teacher evaluations is a sustainable approach to ensuring evaluation rigor. Evaluations will be higher quality where states provide specific tools, models and detailed frameworks for conducting and scoring teacher evaluations.

Even more worrisome are the states in which districts will be left to their own devices, without any attempted oversight. There is good reason to be skeptical that all districts in such states will have the capacity or will to really implement strong teacher evaluation systems.

8. States should start with annual evaluations and modify from there for highly effective teachers once the system is fully operational.

Modifying an evaluation system to allow for less than full fledged annual evaluation for consistently high performers may be sensible in some states, given issues of capacity, but states shouldn't start out that way. States such as **Delaware** that make sure teachers meet student growth expectations each year but require full evaluations for any teacher who fails to meet student achievement expectations may be a reasonable compromise. However, this kind of flexibility must be the result of a good working system, not a starting point for evaluation policies.

9. States and districts should use third party evaluators when possible.

States should encourage districts to use independent, third party evaluators to conduct new teacher evaluations. A third party evaluator can provide important feedback on the evaluation process and important checks for principals and other administrators typically charged with implementing teacher performance reviews. A neutral party who is a demonstrated effective teacher may be able to provide feedback to other teachers on instructional practice in a way that is non-threatening.

10. A scarlet letter isn't appropriate teacher effectiveness policy.

States like **Indiana**, **Michigan** and **Florida** require notification of parents if their child is placed in a classroom with an ineffective teacher. Some think this is good accountability policy. NCTQ thinks this does a tremendous disservice to the teaching profession. If a district has evidence that a teacher is ineffective, state policy should provide the means for the district to take the necessary steps to remove the individual from the classroom, not humiliate the teacher. Reporting on teacher effectiveness data by the state, district and school level is essential. But when it comes to accountability for ineffective teachers, sending a note home to let families know their child's teacher is not so good is no solution at all. **Rhode Island** may have a better alternative. The state makes each district annually certify to the State Commissioner of Education that they have not allowed any student to be taught by an ineffective teacher for more than one year.

11. Teacher evaluation policy should reflect the purpose of helping all teachers improve, not just low-performers.

Many states are only explicit about tying professional development plans to evaluation results if the evaluation results are bad. Good evaluations with meaningful feedback should be useful to all teachers and if done right, should help design professional development plans for all teachers – not just those who receive poor ratings.

12. Anticipate and address the anxieties a new evaluation system creates for teachers, such as how student results will be assigned to individuals for the purpose of student growth and value-added measures of teacher effectiveness.

Teachers, like most people, are afraid of the unknown. States can do more to anticipate fears and diminish tensions over performance-based evaluations. States that have had open and transparent communication as they develop policies, provided detailed information to districts, schools, and teachers about the goals and purposes of the new systems and been clear about how the evaluations will work, will be in a stronger position to implement new systems.

One of the things causing teachers considerable trepidation is the concern on how they will be matched with students. With increased consequences, teachers are understandably worried about students who transfer into their classroom late in the year, students that are on their rosters for administrative purposes but not actually in their classes, and students they are asked to teach but to whom they may not assign grades. To address the concerns and implement reliable data systems that measure teacher effectiveness, states must develop sound definitions of "teacher of record."[33] If student achievement data are to be tied to teacher evaluations, it is essential that a student's information is tagged to the teacher – or teachers – *actually responsible* for that student. Now that some states are moving forward on using data to make decisions of consequence about teaching and learning, such definitions are imperative. Further, states need to develop systems by which teachers are able to verify the students on their rosters. Both **Louisiana** and **New York** now have sophisticated systems for this purpose.

13. Escape clauses need to be shorn up and loopholes closed that may undermine new teacher evaluation systems.

Whether intentional or accidental, loopholes are already visible in some states' evaluation policies that can undermine their intended rigor. Without quick action to shore up these identified weaknesses, states may find themselves disappointed with the results they achieve and/or fighting unnecessary battles.

There is a disconnect in some states between the clear consequences for ineffective teaching spelled out in their evaluation laws and existing – and unamended – laws governing teacher dismissal.

For example, **Colorado**'s teacher evaluation legislation specifically identifies classroom ineffectiveness as grounds for dismissal. Teachers who receive an ineffective rating are clearly eligible for dismissal. However, the state also retains its dismissal law, which does not articulate that classroom performance is grounds for dismissal. Further, the dismissal law provides the same due process for teachers dismissed for any "just cause," not distinguishing dismissal clearly sanctioned through the evaluation system from terminations for reasons commonly associated with license revocation such as a felony and/or morality violations. Failure to bring these statutes into sync may result in preventable legal challenges. **Delaware** and **Louisiana** have similar alignment issues.

Michigan's legislation leaves open to interpretation whether the state has the authority to review and approve or reject district-designed evaluation systems. This ambiguity could slow implementaion timelines if districts challenge the state's authority. In many states, this authority is new territory, only made more difficult to assert if not solidly established.

33 The Center for Educational Leadership and Technology (http://www.celtcorp.com/) has developed a template statement to help states identify teachers of record. A teacher of record is an "educator" who is responsible for a "specified proportion" of a student's "learning activities" that are within a "subject or course" and are aligned to "performance measures".

Other potentially large loopholes include selective provisions which allow collective bargaining agreements to supersede or nullify statutory requirements. In **New York**, for example, nearly every aspect of teacher evaluation is subject to negotiation.

14. States need to get on top of policy plans for equitable distribution of effective teachers now.

A particular goal of improved teacher evaluation systems should be to make sure that high-need schools have the same access to and share of the very best teachers as more affluent schools and districts. The legislative and regulatory policies and provisions we reviewed don't speak to how teacher evaluations should be linked to equitable distribution, but we think states need to address this issue early on in the development of teacher evaluation and effectiveness policy. Without some proactive planning now, the exact opposite of more equitable distribution could occur. Once ratings are issued, there could be a feeding frenzy of schools and districts with more resources trying to lure away the most effective teachers. States and districts need to be planning now with a set of timely policies or incentives to attract and retain the highest-performing teachers to the schools and students most in need of effective teachers.

15. States need to attend to potential bias with systematic checks of their evaluation system; states also need to maintain flexibility to make adjustments to the system as needed.

We are at the beginning of a new policy era about which there is still much to learn. In light of that, states should implement checks to ensure their evaluation systems are fair and reliable. States should analyze and study these issues regularly and systematically. In particular, states should examine the patterns of performance by subject, by measure, and by types of teacher (such as special education teachers) to look for potential red flags for biases in the results. Building in validity checks across subjects and across types of measures will strengthen state and district efforts and increase confidence in and legitimacy of the systems. States also need to build in the ability to modify systems. A lot of research will be generated in the next few years that will inform systems as they develop. Evaluation systems need to be flexible enough to take advantage of what we learn and be able to adjust.

Conclusion

What are the policy implications of an evaluation system that truly measures teacher effectiveness? If done well, and if policymakers act on the results, the consequences are far-reaching and could change much of what is now standard practice in the teaching profession.

A focus on teacher effectiveness can set the foundation for better targeted policies for struggling teachers, higher standards for teacher preparation programs and fair but rigorous policies for replacing persistently ineffective teachers. Compensating teachers based on effectiveness could help attract and retain the best teachers in the profession. A system that cultivates effectiveness will also be crucial to other reform efforts, from implementing new Common Core State Standards and promoting educational equity, to turning around low-performing schools.

Do the policies outlined in this report have a chance of being implemented well? Building better evaluation protocols and stronger operating rules for teacher evaluation systems is only part of what is necessary. The changes required are also a matter of will – on the part of policymakers, school leaders, designated evaluators and teachers themselves – and a sincere commitment to and investment in teacher effectiveness. Even the best evaluation system can be implemented poorly or undermined. So the policies surveyed in this report simply lay the groundwork for teacher effectiveness policies. In order for performance evaluations to become a meaningful part of an effort to build an effective teacher workforce, the evaluations themselves will need to add real value and provide teachers with real benefits by helping them improve their practice.

As enthusiasts for new performance evaluation policies explore the possibilities related to introducing objective measures into teacher evaluations and taking student achievement and growth seriously as a means of assessing teachers, we still must not forget that appraising performance is an activity that involves personal judgment. This is actually a very good thing. We don't want to enslave ourselves in arbitrary ways to testing systems and quantifiable data sets that prohibit reasoned judgment about individual people and their potential as teachers.

What this policy review and early lessons suggest is that performance-based teacher evaluation must be approached in a *measured, realistic and transparent* way. Performance measures are not perfect and good teachers are not the product of formulas. Conducting teacher performance evaluations that focus on the results and the behaviors that matter most will move us towards a system that recognizes and encourages effective instruction and prepares and values highly-effective teachers.

Appendices

Appendix A:

Links to state legislation/regulations and resources on teacher evaluation sytems

Arizona

Arizona Framework for Measuring Teacher Effectiveness:
http://www.azed.gov/highly-qualified-professionals/teacherprincipal-evaluation/
Revised Statute 15-203:
http://www.azleg.gov/FormatDocument.asp?inDoc=/ars/15/00203.htm&Title=15&DocType=ARS

Arkansas

Act 1209:
http://www.arkansashouse.org/bill/2011R/HB2178

Colorado

S.B. 10-191:
http://www.leg.state.co.us/clics/clics2010a/csl.nsf/fsbillcont3/EF2EBB67D47342CF872576A80027B078?open&file=191_enr.pdf
http://www.cde.state.co.us/research/GrowthModel.htm
Review Colorado's draft Model Evaluation System for Principals and Assistant Principals User Guide (September 2011):
http://www.cde.state.co.us/EducatorEffectiveness/index.asp

Delaware

Delaware Administrative Code 14.106A:
http://delcode.delaware.gov/title14/c012/sc07/index.shtml
http://regulations.delaware.gov/AdminCode/title14/100/106A.shtml#TopOfPage
SB 263:
http://legis.delaware.gov/lis/lis145.nsf/vwlegislation/SB+263
See Delaware's Performance Appraisal System Guide for Teachers (August 2011):
http://www.doe.k12.de.us/csa/dpasii/ti/DPASIITeacherFullGuide-9-7-11.pdf

District of Columbia Public Schools (DCPS)

IMPACT:
http://dcps.dc.gov/DCPS/impact
See DCPS's IMPACT Guidebooks:
http://www.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/IMPACT+Guidebooks

Florida

SB 736, amending Florida Statute 1012.34:
http://laws.flrules.org/files/Ch_2011-001.pdf
http://www.fldoe.org/arra/TeacherEvaluationSystems.asp

Idaho

S.B. 1108, "Minimum Statewide Standards":
http://www.sde.idaho.gov/site/teacherEval/

Illinois

Performance Evaluation Reform Act:
http://www.isbe.state.il.us/PEAC/pdf/PA096-0861_SB315.pdf
http://www.isbe.net/PEAC/

Indiana

Indiana Code 20-28-11-3:
http://www.in.gov/legislative/bills/2011/PDF/SE/SE0001.1.pdf
http://www.doe.in.gov/news/documents/May16FAQ_Ver1.pdf
http://www.doe.in.gov/news/2011/05-May/TeacherEffectivenessPilot.html
Indiana's RISE Evaluation and Development System Teacher Effectiveness Rubric (July 2011) is available at:
http://www.doe.in.gov/puttingstudentsfirst/documents/rise_2011-07-10_teacher_effectiveness_rubric_draft.pdf

Louisiana

Act 54:
http://www.act54.org/

Maryland

Education Reform Act of 2010:
http://www.governor.maryland.gov/documents/ERA2010.pdf
COMAR 13a.07.04:
http://www.dsd.state.md.us/comar/SubtitleSearch.aspx?search=13A.07.04.

Massachusetts

603CMR 35:
http://www.doe.mass.edu/lawsregs/603cmr35.html?section=11
http://www.doe.mass.edu/mcas/growth/

Michigan

The Revised School Code Act 451 of 1976 380.1249:
http://www.legislature.mi.gov/(S(2oc5pyjsbpvpx355koez2n45))/mileg.aspx?page=getObject&objectName=mcl-380-1249

Minnesota

HF 26:
https://www.revisor.mn.gov/revisor/pages/search_status/status_detail.php?b=House&f=HF26&ssn=1&y=2011

Nevada

A.B. 229, amending Nevada Revised Statutes 391:
http://www.leg.state.nv.us/Session/76th2011/Bills/AB/AB229_R2.pdf

New York

Regulations:
http://www.regents.nysed.gov/meetings/2011Meetings/May2011/511bra4.pdf

Ohio

HB 153:
http://education.ohio.gov/GD/Templates/Pages/ODE/ODEDetail.aspx?page=3&TopicRelationID=521&ContentID=108217

Oklahoma

Oklahoma Statutes 70-6-101.10, -.16:
http://www.oklegislature.gov/osstatuestitle.html
SB 2033:
http://www.oklegislature.gov/AdvancedSearchForm.aspx

Rhode Island

The RI Model: Guide to Evaluating Building Administrators and Teachers:
http://www.ride.ri.gov/EducatorQuality/EducatorEvaluation/Docs/RIModelGuide.pdf
http://www.ride.ri.gov/commissioner/RaceToTheTop/
http://www.ride.ri.gov/assessment/DOCS/RIGM/RIGM_Pamphlet_FINAL-Spring_2011.pdf
http://www.ride.ri.gov/EducatorQuality/EducatorEvaluation/Docs/TeacherFAQ_01_11_11.pdf

Tennessee

First to the Top:
http://tn.gov/firsttothetop/
http://www.state.tn.us/education/TEAC.shtml

Appendix B

Arizona

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Evaluations occur twice a year for probationary teachers. The timing of evaluations is not specified.
Who is responsible for development of the evaluation system?	Districts are responsible for developing evaluations consistent with state framework.
What measures are required?	All district evaluations must include three components: 1) classroom-level data, 2) school-level data and 3) teaching performance. For teachers with classroom-level data available, additional school-level measure options include aggregate state assessment data, AP/IB results, survey data and other reliable measures. For teachers without reliable classroom data, school-level measures are required.
How is student growth factored?	Evaluations must include quantitative data on student academic progress. These objective data must account for between 33 percent and 50 percent of evaluation measures. Classroom-level data must account for 33 percent and school-level data may account for up to 17 percent for a total of no more than 50 percent of evaluation.
Is student achievement the preponderant criterion?	No.
How many and what are the evaluation categories/ratings?	Not specified.
Who is responsible for conducting evaluations?	Each school district's governing board must designate and ensure qualified teachers to serve as evaluators.
What observation procedures and feedback are specified?	Multiple observations are required.
What are the required uses for evaluation results?	Each evaluation must include recommendations as to areas of improvement, if the performance of the teacher warrants improvement. A board designee must confer with the teacher to make specific recommendations. Assistance and opportunities must be provided for the certificated teacher to improve performance. After a reasonable period of time, the designee must follow up with the teacher to ascertain whether that teacher is demonstrating adequate classroom performance.
What employment consequences are tied to evaluation results?	State policy authorizes districts to develop definitions of inadequate classroom performance. Districts have discretion to propose what actions, if any, the school district plans to take regarding dismissal of teachers based on inadequate classroom performance. The state's pay for performance program requires that the placement of teachers on the career ladder be based on more than one measure of performance that includes increasingly higher levels of student academic progress, the use of various methods of progress assessments by local districts and procedures for review of student progress. The restructured salary schedule must be based on performance and not on experience and education.
What is the state's implementation timeline?	By school year 2012-2013, districts must use an evaluation instrument that meets state requirements for annual evaluations of all teachers.

Colorado

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. The timing of observations/evaluations is not specified.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model with district opt-in.
What measures are required?	All district evaluations must measure five professional practices: 1) know content, 2) establish environment, 3) facilitate learning, 4) reflect on practice and 5) demonstrate leadership and student growth.
How is student growth factored?	50 percent of evaluation rating is based on: 1) measures of individually-attributed growth; 2) a measure of collectively-attributed growth whether on a school-wide basis or across grades or subjects; 3) when available, statewide summative assessment results; and 4) Colorado Growth model for subjects with annual statewide summative assessment results available in two consecutive grades.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Three performance levels: highly effective, effective and ineffective. While Colorado's legislation specifies three categories, draft regulations indicate that there will be four.
Who is responsible for conducting evaluations?	Each principal or administrator who is responsible for evaluating licensed personnel shall keep records and documentation for each evaluation conducted. Each principal and administrator who is responsible for evaluating licensed personnel shall be evaluated as to how well he or she complies with the school district's evaluation system.
What observation procedures and feedback are specified?	Probationary teachers must receive at least two documented observations and one evaluation that result in a written evaluation report each academic year. Beginning with the 2012-2013 school year, all other teachers must receive a written evaluation report each academic year.
What are the required uses for evaluation results?	Each teacher must be provided with an opportunity to improve effectiveness through a teacher development plan. For nonprobationary teachers, a remediation plan must be developed by the district and must include professional development opportunities. The teacher must be given a reasonable period of time to remediate deficiencies. If the next evaluation shows effective performance, no further action must be taken. Otherwise, the evaluator must either make additional recommendations for improvement or may recommend dismissal.
What employment consequences are tied to evaluation results?	Probationary teachers must earn three consecutive effective ratings to earn the equivalent of tenure. Veteran, or non-probationary, teachers who receive two consecutive ineffective ratings return to probationary status and have a year to improve or face termination.
What is the state's implementation timeline?	In March 2011 the State Council for Educator Effectiveness provided the state board with recommendations on evaluation methods. In 2011-2012: State education department must work with school districts to develop performance evaluation systems. By 2012-2013: New systems must be implemented and tested. By 2013-2014: Statewide implementation and evaluations shall be considered in acquisition of non-probationary status. In 2014-2015: Evaluation system finalized statewide.

Delaware

What is the frequency and timing of teacher evaluations?	New teachers must receive one summative evaluation rating every year. Experienced teachers who earn a rating of highly effective must receive a summative evaluation at least once every two years. However, the student improvement evaluation component must be evaluated every year, and teachers cannot be rated effective unless they have met growth targets. If a highly effective teacher does not achieve a satisfactory rating on the student improvement component, the teacher must receive a summative evaluation the following year.
Who is responsible for development of the evaluation system?	State has a single statewide teacher evaluation system (DPAS II). Local districts are permitted to use their own evaluation instruments in addition to the statewide system, but not in place of statewide instruments.
What measures are required?	The state model measures: 1) planning and preparation, 2) classroom environment, 3) instruction, 4) professional responsibilities and 5) student improvement. For tested grades and subjects, student improvement measures are based on scores on the statewide assessment, and other measures of student learning that are rigorous and comparable across classrooms. For nontested grades and subjects, the state requires alternative measures of student learning such as scores on pre-tests and end-of-course tests, performance on English language proficiency assessments, and other measures that are rigorous and comparable across classrooms to be included in the evaluations. Districts can develop and implement other assessment tools that measure student improvement, including assessments in other content areas, with state approval. Districts also may implement evaluation measures in addition to the state system requirements. Among other options for evidence are: classroom management documents; examples of student work/assignments; and communication logs with parents.
How is student growth factored?	Teacher cannot be rated effective overall if the student growth expectations for the teacher's students are not met. Schoolwide assessment measures account for 30 percent of student improvement component and student cohort assessment measures account for 20 percent. Teacher specific assessment measures account for 50 percent of student improvement component.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance levels: highly effective, effective, needs improvement and ineffective.
Who is responsible for conducting evaluations?	Credentialed evaluators are usually the supervisor of the teacher, who must complete training and receive a certificate of completion, which is valid for five years and is renewable upon completion of professional development.
What observation procedures and feedback are specified?	New teachers must receive a minimum of two announced observations and one unannounced observation every year. Experienced teachers who earn a rating of highly effective on their most recent summative evaluation must receive a minimum of one announced observation each year. After each observation, the teacher and evaluator participate in a post-observation conference to discuss the teacher's performance.
What are the required uses for evaluation results?	Teachers who receive an overall rating of needs improvement or ineffective on the summative evaluation, or a rating of unsatisfactory on any appraisal component regardless of the overall rating, must be put on an improvement plan.
What employment consequences are tied to evaluation results?	The state requires that teachers must show two years of satisfactory student growth (evidenced by satisfactory ratings in the student improvement component of the teacher appraisal process) within a three-year period before they receive tenure. Teachers with two consecutive years of ineffective ratings or who earn a combination of ineffective and unsatisfactory ratings for three consecutive years are considered to have patterns of ineffective teaching and are eligible for dismissal.
What is the state's implementation timeline?	The state's regulations regarding teacher evaluations went into effect July 1, 2011.

District of Columbia Public Schools (DCPS)

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. The first administrator observation must occur between September 21 and December 1, and the first master educator observation must occur between September 21 and February 1.
Who is responsible for development of the evaluation system?	DCPS has a single districtwide teacher evaluation system (IMPACT).
What measures are required?	The District's system measures: 1) student achievement (both individual and school value added), 2) instructional expertise, 3) commitment to school community and 4) professionalism. Individual and school level value-added data are included for teachers in grades 4-8. Teacher-assessed data on tests other than statewide assessments and school level value-added data are included for all other teachers.
How is student growth factored?	DCPS's system, IMPACT, requires that 50 percent of the evaluation score be based on the teacher's impact on students' achievement; 5 percent on schoolwide value-added measures.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance levels: highly effective, effective, minimally effective and ineffective.
Who is responsible for conducting evaluations?	Three of the observations are conducted by an administrator, and two are conducted by an impartial third-party called a master educator.
What observation procedures and feedback are specified?	DCPS teachers are formally observed five times annually. Within 15 days of each observation, the observer must meet with the teacher to share ratings, provide feedback, and discuss next steps for professional growth.
What are the required uses for evaluation results?	Those who are rated minimally effective are encouraged to take advantage of professional development opportunities provided by DCPS. They are held at current salary step until they earn a rating of effective or higher.
What employment consequences are tied to evaluation results?	Those teachers who receive a minimally effective rating for two consecutive years are subject to separation from the school system. Teachers who receive an ineffective rating are subject to separation from the school system. Members of the Washington Teachers' Union who receive highly effective ratings are eligible for additional compensation under new contract.
What is the state's implementation timeline?	IMPACT was implemented in 2009-10. A small number of changes were made to the system for 2010-11. DCPS also instituted a process for principals to contest a teacher dismissal based on IMPACT.

Florida

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. New teachers must be evaluated at least twice in the first year of teaching. The timing of observations/evaluations is not specified.
Who is responsible for development of the evaluation system?	Districts design systems consistent with state framework and subject to state review and approval.
What measures are required?	Districts design but state promotes model with student growth requirements plus four domains: 1) classroom strategies and behaviors, 2) planning and preparing, 3) reflections on teaching and 4) collegiality and professionalism, in addition to student growth requirements. Growth requirements include state-adopted student growth measures for grades and subjects with state assessment data, district assessments for subjects and grades not covered by state tests, and principal-set targets where no district tests are available. System must include at least one additional metric of student performance to have a multi-metric evaluation for teachers in the year before a "milestone event" and parents must have an opportunity for input on teacher performance ratings.
How is student growth factored?	At least 50 percent of the evaluation must be based on data and indicators of student learning growth assessed annually by statewide assessments. The student learning growth portion of the evaluation must include growth data for students assigned to the teacher over the course of at least three years. If three years of data are not available, the years for which data are available must be used and the percentage of the evaluation based upon student learning growth may be reduced to not less than 40 percent.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance levels: highly effective, effective, needs improvement and unsatisfactory.
Who is responsible for conducting evaluations?	The individual responsible for supervising the employee must evaluate the employee's performance.
What observation procedures and feedback are specified?	New teachers must be observed at least twice in the first year of teaching. Other teacher must be observed at least once a year.
What are the required uses for evaluation results?	If a teacher receives an unsatisfactory evaluation, the evaluator must make recommendations as to specific areas of unsatisfactory performance, and provide assistance in helping to correct deficiencies within a prescribed period of time. A teacher with unsatisfactory performance is then placed on performance probation for 90 days. During these 90 days, the teacher must be evaluated periodically and apprised of progress achieved, and must be provided assistance and in-service training opportunities to help correct deficiencies. Within 14 days after 90-day period, evaluator re-evaluates to see whether deficiencies have been corrected, and whether teacher should continue employment or be terminated.
What employment consequences are tied to evaluation results?	Beginning July 1, 2011, an annual contract may be awarded to a teacher who has completed a probationary contract, if that teacher: 1) has been recommended by the district school superintendent for the annual contract based upon the individual's evaluation and 2) has not received two consecutive annual performance evaluation ratings of unsatisfactory, two annual performance evaluation ratings of unsatisfactory within a 3-year period, or three consecutive annual performance evaluation ratings of needs improvement or a combination of needs improvement and unsatisfactory.
What is the state's implementation timeline?	Beginning in the 2011-2012 school year, each school district shall measure student-learning growth using a formula approved by the State Commissioner.

Idaho

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Evaluations are to occur twice a year for new/probationary teachers. The first of the evaluations for probationary teachers must occur before January 1.
Who is responsible for development of the evaluation system?	Districts design systems consistent with state framework and subject to state review and approval.
What measures are required?	In addition to student growth, the state requires four domains to be assessed in evaluations: 1) planning and preparation, 2) learning environment, 3) instruction and use of assessment and 4) professional responsibilities. The state requires objective measures of student growth, as determined by the board of trustees. Evaluations must include input from parents and guardians of students as a factor.
How is student growth factored?	At least 50 percent of the evaluation shall be based on objective measure(s) of growth in student achievement.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Two performance levels: minimum of two categories to address proficient and unsatisfactory.
Who is responsible for conducting evaluations?	Administrators; districts must ensure that all administrators responsible for performing evaluations be trained in the district approved evaluation model.
What observation procedures and feed-back are specified?	Observations are required but frequency is not specified.
What are the required uses for evaluation results?	Not specified. Districts can propose procedures to provide remediation in those instances where remediation is determined to be an appropriate course of action.
What employment consequences are tied to evaluation results?	Not specified. Districts have discretion to propose what actions, if any, the school district plans to take as a result of evaluations.
What is the state's implementation timeline?	By Fall 2011, districts are expected to begin full implementation of the teacher evaluation model. By February, 2013 and every February following, the first half of a teacher's evaluation is due. This half of a teacher's evaluation is based on the Charlotte Danielson Framework and includes the parent/guardian input component. By the end of the 2013 school year and subsequent years the second half of a teacher's evaluation is due. This half of a teacher's evaluation is based on student achievement as determined by the local school board.

Illinois

What is the frequency and timing of teacher evaluations?	The state requires that probationary teachers be evaluated annually and that nonprobationary teachers be evaluated at least once every two years. However, any teacher in contractual continued service whose performance is rated as either needs improvement or unsatisfactory must be evaluated at least once in the school year following the receipt of such rating. The timing of evaluations is not specified.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model with district opt-in. Illinois has stipulated that if a district's committee cannot reach an agreement within 180 days, then the district must implement the state's model evaluation plan.
What measures are required?	Components to be determined by September 2012. State law requires that system must consider: the teacher's attendance, planning and instructional methods; classroom management, where relevant; and competency in the subject matter taught.
How is student growth factored?	The state requires that the use of data and indicators of student growth must be "significant" factors in teacher evaluations.
Is student achievement the preponderant criterion?	No.
How many and what are the evaluation categories/ratings?	Four performance levels: excellent, proficient, needs improvement and unsatisfactory.
Who is responsible for conducting evaluations?	Evaluators are qualified administrators who complete a pre-qualification program. The program must involve rigorous training and an independent observer's determination that the evaluator's ratings properly align to the requirements established by the State Board.
What observation procedures and feedback are specified?	Personal classroom observations are required but frequency is not specified.
What are the required uses for evaluation results?	Those who receive a rating of needs improvement must be placed on professional development plan to address those areas. Those rated unsatisfactory must be placed on remediation plan. If a teacher does not improve on the remediation plan, with a rating equal to or better than satisfactory or proficient, then he/she is subject to dismissal.
What employment consequences are tied to evaluation results?	Teachers will no longer automatically receive tenure after four years in the classroom regardless of performance. Tenure decisions will now be based on performance evaluations by requiring teachers to earn two proficient or excellent ratings in years two through four of probationary period, with a proficient or excellent rating in fourth year. Illinois specifically identifies classroom ineffectiveness as grounds for dismissal. For teachers placed on remediation plans for poor performance that receive a subsequent unsatisfactory performance rating within three years, the school district may forego remediation and seek dismissal.
What is the state's implementation timeline?	By September 1, 2012, student growth must be significant factor in teacher evaluations.

Indiana

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Existing regulations state that new teachers (non-permanent and semi-permanent") must be formally evaluated before December 31. If requested by the teacher, an additional evaluation may be scheduled on or before March 1 of the following year.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model with district opt-in.
What measures are required?	In addition to a core professionalism rubric, the state's model rubric includes three domains: 1) purposeful planning, 2) effective instruction and 3) teacher leadership. Student achievement and growth measures are based on assessment results from: 1) statewide assessments; and 2) methods for assessing growth for teachers in areas not measured statewide, including results from locally developed assessments and other tests.
How is student growth factored?	Objective measures of student achievement and growth must "significantly inform" the evaluation.
Is student achievement the preponderant criterion?	No.
How many and what are the evaluation categories/ratings?	Four performance levels called for by Indiana Teacher Effectiveness Pilot: highly effective, effective, improvement necessary and ineffective.
Who is responsible for conducting evaluations?	Individuals with demonstrated records of effective teaching and principal approval to conduct evaluations.
What observation procedures and feed-back are specified?	Observations are required but frequency is not specified.
What are the required uses for evaluation results?	Not specified.
What employment consequences are tied to evaluation results?	A contract with an established teacher may be cancelled if the teacher receives two consecutive ineffective ratings or if the teacher receives an ineffective or improvement necessary rating in three years of any five year period. Raises cannot be given to teachers who are not evaluated effective or highly effective. Starting in July 2012, Indiana requires local salary scales to be based upon a combination of factors. Years of teacher experience and content area degrees beyond the requirements for employment may not account for more than thirty-three percent of the calculation. The remaining calculation is based on results of the teacher evaluation based on a number of factors including teacher performance and student achievement, which should include but not be limited to test results.
What is the state's implementation timeline?	The state evaluation rubric (RISE) and alternative models will be piloted in the fall of 2011. Information from the pilot will be used to further fine-tune the model.

Louisiana

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. The timing of the evaluation is not specified.
Who is responsible for development of the evaluation system?	State has a single statewide teacher evaluation system (COMPASS).
What measures are required?	Half of every educator's evaluation rating centers on the growth their students make over the course of the school year captured by the GO-Index (Growth Outcomes Index). The remaining half, the SITE-Index, is based on traditional evaluation techniques, such as classroom observations. Outcome measures will be based on results from value-added assessment model determined by state board for grades where data are available. For grades, subjects and personnel for which value-added data are not available, the state board shall establish growth measures.
How is student growth factored?	The state' requires that 50 percent of evaluations be based on evidence of growth in student achievement using a value-added assessment model. In Louisiana, value-added assessment begins by using each student's history of test score in core subjects for up to three years, and critical individual student factors (such as special education disability diagnosis, attendance, discipline history, and free lunch status) to estimate a student's expected level of achievement for the current year. This estimate is derived from each individual student's data, and the data for all other test takers in Louisiana. In order for a student's assessment results to contribute to a value-added assessment for that teacher, the student must: have been enrolled in that school from early fall until testing time; have a prior year of standardized test data; and take the regular state assessments. The assessment compares actual student achievement for eligible students to the predicted achievement to determine if the student has made more, less, or a typical amount of progress. The results for all students in a teacher's assignment are then combined for that teacher.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Two performance levels: Effective and ineffective. While state legislation only identifies two performance categories, state evaluation materials suggest there will be additional categories.
Who is responsible for conducting evaluations?	Evaluators are defined as school principal or vice principal, or respective supervisory level designees.
What observation procedures and feedback are specified?	Observation is required, along with a post-observation conference to discuss commendation and recommendations. Teachers must be provided a copy of the evaluation and the evaluators' data recording forms.
What are the required uses for evaluation results?	Any teacher not deemed effective will be placed in an intensive assistance program and then must be formally re-evaluated. Program must include an expected time line for achieving objectives, which must not exceed two years. If teacher is still ineffective, then board will initiate termination proceedings.
What employment consequences are tied to evaluation results?	If a teacher's evaluation demonstrates that he has met the standard for effectiveness as determined by the board, using value-added data, for three years during the initial certification or renewal process, a certificate shall be issued or renewed unless the board receives evidence from the local board, through an appeal. Teachers who receive an ineffective rating three or more times during a certification cycle will not be recertified. The board shall determine a standard for highly effective teachers for use by local boards to recognize, reward, and retain teachers who demonstrate a high level of effectiveness.
What is the state's implementation timeline?	By 2012-13, the state's value-added model will be implemented.

Maryland

What is the frequency and timing of teacher evaluations?	Most teachers must be evaluated annually. Probationary teachers must be formally evaluated at least once a semester. Standard Professional Certificate holders are to be evaluated annually. Advanced Professional Certificate holders must receive an evaluation at least twice during the validity period of the certificate, with the first evaluation occurring during the initial year. Timing of evaluations is not specified.
Who is responsible for development of the evaluation system?	Districts design systems consistent with state framework and subject to state review and approval.
What measures are required?	Evaluation standards are focused on student growth and professional practice. State framework on professional practice requires evaluation of: 1) planning and preparation, 2) instruction, 3) classroom environment and 4) professional responsibilities. Districts can propose other additional local priorities and measures with state approval such as scholarship, management skills, professional ethics and interpersonal relationships. An existing or newly created single assessment may not be used solely as evidence of student growth. Districts choose measures based on a menu of approved options. If a statewide assessment is available for a teacher, it must be one of the measures used.
How is student growth factored?	Multiple measures of achievement total 50 percent of evaluation. No single measure can account for more than 35 percent.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Three performance levels: Maryland's pilot evaluation program is using highly effective, effective and ineffective.
Who is responsible for conducting evaluations?	Trained evaluator.
What observation procedures and feedback are specified?	Maryland requires that evaluators conduct at least two observations during the school year. State regulations require that a written observation report must be shared with teacher within a reasonable period of time.
What are the required uses for evaluation results?	An unsatisfactory evaluation must include at least one observation by someone other than the immediate supervisor. Teachers may appeal overall ratings of unsatisfactory; the burden of proof is on teacher.
What employment consequences are tied to evaluation results?	None specified.
What is the state's implementation timeline?	Districts must implement new evaluations by 2012-13 as a no-fault system that year, with full implementation during the 2013-14 school year.

Michigan

What is the frequency and timing of teacher evaluations?	Most teachers must be evaluated annually. If a teacher is rated as highly effective on three consecutive annual year-end evaluations, the school district, intermediate school district, or public school academy may choose to conduct a year-end evaluation every other year. Timing of observations/evaluations is not specified.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model is presumptive model for each district in the state although districts are permitted to propose alternative approaches that meet the same standards as the state-designed system.
What measures are required?	Student growth is specified, other standards to be determined by governor's council on educator effectiveness. System may include the following: instructional leadership abilities, teacher and pupil attendance, professional contributions, training, progress report achievement, school improvement plan progress, peer input, and pupil and parent feedback. System must include specific performance goals that will assist in improving effectiveness for the next school year.
How is student growth factored?	For the annual year-end evaluation for the 2013-2014 school year, at least 25 percent will be based on student growth and assessment data. For the annual year-end evaluation for the 2014-2015 school year, at least 40 percent shall be based on student growth and assessment data. Beginning with the annual year-end evaluation for the 2015-2016 school year, at least 50 percent will be based on student growth and assessment data.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Multiple rating categories required. Required number and names of performance levels are not yet specified.
Who is responsible for conducting evaluations?	School administrator or designee.
What observation procedures and feedback are specified?	Multiple observations are required unless teacher is rated effective or highly effective.
What are the required uses for evaluation results?	Teacher must be given "ample opportunities for improvement."
What employment consequences are tied to evaluation results?	Michigan bases tenure on evidence of classroom effectiveness. Michigan's has increased the probationary period for new teachers to five years. The state also articulates that a teacher has not successfully completed this probationary period unless he or she has been rated as effective or highly effective on the three most recent annual performance evaluations. Michigan also identifies classroom ineffectiveness as grounds for dismissal. If a teacher is rated as ineffective on three consecutive annual year-end evaluations, the district shall dismiss the teacher.
What is the state's implementation timeline?	Not later than September 1, 2011, the board of a school district or intermediate school district or board of directors of a public school academy shall adopt and implement for all teachers and school administrators a rigorous, transparent, and fair performance evaluation system.

Minnesota

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Evaluations are to occur three times a year for probationary teachers. The first of the three evaluations for probationary teachers must occur within 90 days of the beginning of teaching services.
Who is responsible for development of the evaluation system?	Districts are responsible for developing evaluations consistent with state framework.
What measures are required?	Only value added or student growth measures are specified by state. Evaluation system must be based on professional teaching standards. Local school board and teacher representatives agree on further detail. System elements must include option for teachers to present a portfolio demonstrating professional growth and teachers' own performance assessments based on student work samples. System must use longitudinal data on student engagement and connection and other student outcome measures aligned with curriculum.
How is student growth factored?	As a basis for 35 percent of teacher evaluation results school boards must: 1) use an agreed upon teacher value-added assessment model for grades and subjects where data are available and 2) establish state or local student growth measures where value-added data are not available.
Is student achievement the preponderant criterion?	No.
How many and what are the evaluation categories/ratings?	Not specified.
Who is responsible for conducting evaluations?	State requires trained and qualified evaluators, such as school administrators, to conduct peer reviews and evaluations.
What observation procedures and feed-back are specified?	Not specified.
What are the required uses for evaluation results?	Districts must give teachers not meeting professional teaching standards support to improve through a teacher improvement process that includes established goals and timelines.
What employment consequences are tied to evaluation results?	State requires that districts discipline a teacher for not making adequate progress in the teacher improvement process.
What is the state's implementation timeline?	The state department of education must create and publish a teacher evaluation process that complies with the requirements for an annual teacher evaluation and peer review process to be implemented beginning in the 2014-2015 school year.

Nevada

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. For probationary teachers, a conference and a written evaluation must be concluded not later than December 1; February 1; and April 1 of each school year of the probationary period.
Who is responsible for development of the evaluation system?	Districts are responsible for developing evaluations consistent with state framework. A copy of the evaluation procedure must be filed with the state.
What measures are required?	State specifies precisely only student achievement requirement. Evaluations must include the following: an evaluation of classroom management skills, a review of lesson plans or grade book, an evaluation of whether the curriculum taught is aligned with standards, and an evaluation of whether the teacher is appropriately addressing the needs of students.
How is student growth factored?	Information on pupil achievement maintained by the automated system of accountability information for the state must account for at least 50 percent of teacher evaluations.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance categories: highly effective, effective, minimally effective and ineffective.
Who is responsible for conducting evaluations?	Administrators.
What observation procedures and feedback are specified?	Observations totaling not less than 60 minutes per evaluation period are required, with one observation of at least 30 minutes. For probationary teachers, a conference and a written evaluation must be concluded not later than December 1; February 1; and April 1 of each school year of the probationary period. Other teachers must receive a copy of each evaluation within 15 days.
What are the required uses for evaluation results?	Districts are responsible to ensure that a "reasonable effort" is made to correct deficiencies upon the request of a teacher who needs assistance.
What employment consequences are tied to evaluation results?	A probationary teacher who completes a 3-year probationary period and receives a satisfactory designation on each performance evaluation for two consecutive years is eligible for post-probation status. Teacher ineffectiveness is grounds for dismissal. All post-probationary teachers will return to probationary status if they receive two consecutive years of unsatisfactory evaluations. If the board of trustees of a school district determines that a reduction in the existing workforce of the licensed educational personnel in the school district is necessary, the decision to lay off a teacher or an administrator must not be based solely on the seniority of the teacher or administrator and may include, without limitation, a consideration of, among other factors, results of teacher evaluations.
What is the state's implementation timeline?	Commencing with the 2013-2014 school year, districts will implement and carry out the policies for evaluations of teachers.

New York

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Timing of evaluation is not specified.
Who is responsible for development of the evaluation system?	Districts are responsible for developing evaluations consistent with state framework.
What measures are required?	Evaluation system includes statewide student growth measures; locally selected measures of student achievement; teacher observations, school visits and other measures to provide teachers with detailed, structured feedback on professional practice. Menu of state-approved rubrics for assessing New York State Teaching Standards is to be determined. Optional additional element includes structured reviews of student work, portfolios, feedback from students, parents or others using structured surveys and teacher self-reflection.
How is student growth factored?	40 percent of teacher evaluations are based on student achievement. A total of 20 percent is based on student growth on state assessments or a comparable measure of student achievement growth (increases to 25 percent upon implementation of a value-added growth model); 20 percent is based on locally-selected measures of student achievement that are determined to be rigorous and comparable across classrooms (decreases to 15 percent upon implementation of a value-added growth model).
Is student achievement the preponderant criterion?	No.
How many and what are the evaluation categories/ratings?	Four performance levels: highly effective, effective, developing and ineffective.
Who is responsible for conducting evaluations?	To the extent practicable, the building principal, or his or her designee, is to be the lead evaluator of a classroom teacher.
What observation procedures and feedback are specified?	Evaluations must include one or more classroom observations, which account for half of the 60 percent of evaluation score not included in student achievement measures.
What are the required uses for evaluation results?	If a teacher is rated developing or ineffective, the school district is required to develop and implement a teacher or principal improvement plan.
What employment consequences are tied to evaluation results?	New York has a three-year probationary period for new teachers. At the conclusion of this period, the state's policy regarding tenure decisions requires evaluation of the "candidate's effectiveness over the applicable probationary period in contributing to the successful academic performance of his or her students." Tenured teachers and principals with a pattern of ineffective teaching or performance, defined as two consecutive annual ineffective ratings, may be charged with incompetence and considered for termination through an expedited hearing process.
What is the state's implementation timeline?	The Department recommends that, to the extent possible, districts begin the process of rolling this system out for the evaluation of all classroom teachers and building principals in the 2011-2012 school year.

Ohio

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. New teachers must be evaluated twice a year. For veteran teachers evaluations must be completed by April 1; teachers must receive written reports by April 10. For new teachers the first evaluation must be completed by January 15, with the written report submitted to the teacher by January 25; the second must be scheduled between February 10 and April 1, with the written report submitted to the teacher by April 10.
Who is responsible for development of the evaluation system?	Districts are responsible for developing evaluations consistent with state framework.
What measures are required?	Ohio Teacher Evaluation System includes, in addition to student academic growth requirements, measures of: 1) professional goal-setting, 2) formative assessment of teacher performance and communication and 3) professionalism. Evaluation includes a self-assessment for all evaluation components. Experienced teachers who perform at an accomplished level may choose to complete a professional project as part of the annual evaluation.
How is student growth factored?	State requires that 50 percent of evaluations be based on student learning measures that include performance on the statewide achievement tests and the college and work ready assessments, as well as the value-added progress dimension.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance categories: accomplished, proficient, developing and ineffective.
Who is responsible for conducting evaluations?	Each evaluation must be conducted by one or more of the following: 1) A person who holds a license designated for being a superintendent, assistant superintendent, or principal; 2) A person who holds a license designated for being a vocational director or a supervisor; 3) A person designated to conduct evaluations under an agreement providing for peer review entered into by the board and representatives of teachers employed by the board.
What observation procedures and feedback are specified?	Evaluations must include at least two observations. State requires that each teacher be provided with a written report of the results.
What are the required uses for evaluation results?	Each teacher must be provided with a written report of the results of the teacher's evaluation that includes specific recommendations for any improvements needed in the teacher's performance, suggestions for professional development that will enhance future performance in areas that do not meet expected performance levels, and information on how to obtain assistance in making needed improvements.
What employment consequences are tied to evaluation results?	Due to the pending referendum on its education reform bill, Ohio currently has two versions of its education code pertaining to teacher evaluation on its books. One version specifically addresses removal of poorly performing teachers, but the other does not.
What is the state's implementation timeline?	By April 30, 2012, the superintendent of public instruction must submit to the Board an evaluation framework. The state board shall adopt the recommended or modified frameworks not later than July 1, 2012. Not later than July 1, 2013, districts must adopt evaluation policy based on state framework.

Oklahoma

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Existing regulations stipulate that probationary teachers must be evaluated at least two times per school year, once prior to November 15 and once prior to February 10 of each year.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model with district opt-in (according to September 2011 recommendations posted for public comment).
What measures are required?	In addition to student achievement, the state is considering three framework options for teacher evaluation: 1) Danielson's Framework for Teaching; 2) Marzano's Causal Teacher Evaluation Model; and 3) Tulsa's TLE Observation and Evaluation System. State requires measures of academic growth using multiple years of standardized test data. Where there is no state-mandated testing measure, evaluations must include objective measures including student performance on unit or end-of-year tests and overall school growth.
How is student growth factored?	50 percent of the ratings of teachers are to be based on quantitative components divided as follows: 35 percentage points based on student academic growth using multiple years of standardized test data, as available, and 15 percentage points based on other academic measurements.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Five performance levels: superior, highly effective, effective, needs improvement and ineffective.
Who is responsible for conducting evaluations?	All certified personnel must be evaluated by a principal, assistant principal or other trained certified individual designated by district board.
What observation procedures and feedback are specified?	Evaluations must include "observable and measurable characteristics of personnel and classroom practice"; frequency not specified.
What are the required uses for evaluation results?	All teachers who receive ratings of needs improvement or ineffective must be placed on comprehensive remediation plans and provided with instructional coaching.
What employment consequences are tied to evaluation results?	Tenured teachers rated ineffective for two consecutive years will be terminated; those rated as needs improvement for three years will be terminated; and those who do not average at least an effective rating over a five-year period will be terminated.
What is the state's implementation timeline?	District evaluation policies must be revised by 2013-14 school year to reflect state requirements.

Rhode Island

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. Throughout the year, evaluators will observe teachers, both during longer, announced observations and unannounced observations, which may be shorter. Each teacher should receive at least one long, announced observation in the first semester of the year.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model is presumptive model for each district in the state although districts are permitted to propose alternative approaches that meet the same standards as the state-designed system.
What measures are required?	State model includes evaluation of: 1) student learning, 2) professional practice and 3) professional responsibilities. Student learning will be measured in two ways. Administrators and teachers in each school will work together to set specific, measureable student learning objectives for each grade and subject at the beginning of the year. These student-learning objectives should be standards-based and tailored to reflect the unique learning needs of students in each school. The objectives may be adjusted at the mid-year conference, based on available evidence, to ensure that they remain appropriate. Objectives will be assessed at the end of the year. In addition, starting in school year 2012-2013, teachers who teach reading and mathematics in grades 3 through 7 will receive a score based on the Rhode Island Growth Model – a statistical model that measures how each teacher's students progressed in comparison to students throughout the state with the same score history.
How is student growth factored?	Student learning is predominant component through matrix scoring model.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance levels: highly effective, effective, developing and ineffective.
Who is responsible for conducting evaluations?	In most instances, principals or assistant principals will be evaluators. In some districts and schools, department chairs or other instructional leaders may serve as evaluators. Districts may also choose to use complementary evaluators, who can assist principals in conducting some classroom observations and providing feedback.
What observation procedures and feedback are specified?	Throughout the year, evaluators will observe teachers, both during longer, announced observations and unannounced observations, which may be shorter. A teacher should receive at least four total observations, including both longer, announced and shorter, unannounced observations. Each teacher should receive at least one long, announced observation in the first semester of the year, prior to a mid-year conference. Written feedback should be provided to the teacher within two to three school days of the observation.
What are the required uses for evaluation results?	Teachers who are rated as developing or ineffective at the end of the year will be placed on an individual development plan and will work with an improvement team to assist them with their development over the course of the following year. An improvement team may consist solely of an educator's evaluator, or of multiple people, depending on the educator's needs and the school and district context. The teacher's district will identify personnel actions that may occur if he or she does not adequately improve his or her performance.
What employment consequences are tied to evaluation results?	Rhode Island has discontinued its policy of automatic tenure. Teachers who receive two years of ineffective evaluations will be dismissed. Any teacher with five years of ineffective ratings would not be eligible to have his or her certification renewed by the state. Districts may no longer make teacher assignments that are based solely on seniority.
What is the state's implementation timeline?	The Rhode Island Model Educator Evaluation System will begin to be gradually implemented statewide in the 2011-2012 school year. During the first year of implementation, ratings will be used for development purposes only, and the model will be analyzed and improved based on the experience and feedback of teachers. In school year 2012-2013, districts will implement the full version of the Rhode Island Model, which will incorporate lessons learned from the first year of implementation.

Tennessee

What is the frequency and timing of teacher evaluations?	Annual evaluations are required for all teachers. All teachers on professional licenses will have two observations occurring in each semester. Apprentice teachers will be observed three times in each semester.
Who is responsible for development of the evaluation system?	State designed teacher evaluation model is presumptive model for each district in the state although districts are permitted to propose alternative approaches that meet the same standards as the state-designed system. Districts may apply for State Board approval to use their own instruments.
What measures are required?	In addition to student growth, state uses the TAP rubric, an observation tool that includes 26 indicators of teaching skills focused on: 1) planning, 2) environment, 3) professionalism and 4) instruction. Mandatory criteria include review of prior evaluations and classroom observations. Evaluations based on student growth data from the Tennessee Value-Added Assessment System (TVAAS)or other comparable measure of growth.
How is student growth factored?	50 percent of evaluation is based on student achievement, including 35 percent based on student achievement growth and 15 percent based on other achievement measures.
Is student achievement the preponderant criterion?	Yes.
How many and what are the evaluation categories/ratings?	Four performance levels: advanced, proficient, developing and unsatisfactory.
Who is responsible for conducting evaluations?	Principals have primary responsibility for evaluations. All observers will be trained directly by expert trainers in four-day training sessions across the state. At the end of the four days, observers will be required to pass a certification test.
What observation procedures and feedback are specified?	All teachers on professional licenses will be observed four times annually, with two observations occurring in each semester and at least half of all observations unannounced. Apprentice teachers will be observed six times annually, three in each semester and at least half unannounced. Conferences are required.
What are the required uses for evaluation results?	Not addressed.
What employment consequences are tied to evaluation results?	The evaluations shall be a factor in employment decisions, including but not necessarily limited to promotion, retention, termination, compensation and the attainment of tenure status. New tenure regulations add a requirement for tenure eligibility that the teacher must have received evaluations demonstrating an acceptable level of performance as provided in the evaluation guidelines adopted by the state board of education, during the last two years of their probationary period. Tennessee explicitly makes teacher ineffectiveness grounds for dismissal. Tennessee specifies that tenured teachers who receive two consecutive years of unacceptable performance ratings are returned to probationary status, making them eligible for dismissal.
What is the state's implementation timeline?	The policies shall be implemented prior to the 2011-2012 academic year.

National Council on Teacher Quality

1420 New York Avenue, Suite 800
Washington, D.C. 20005
Tel: 202 393-0020 Fax: 202 393-0095
Web: www.nctq.org

The National Council on Teacher Quality advocates for reforms in a broad range of teacher policies at the federal, state and local levels in order to increase the number of effective teachers.

Subscribe to NCTQ's free monthly electronic newsletter, *Teacher Quality Bulletin*, (www.nctq.org/p/tqb/subscribe.jsp), to stay abreast of trends in federal, state and local teacher policies and the events that help to shape them.

Subscribe to NCTQ's blog, *PDQ*.

Follow us on