

Center for Public Education



TRENDS IN TEACHER EVALUATION

How states are measuring
teacher performance

Jim Hull
Senior Policy Analyst, Center for Public Education
October 2013



Trends in Teacher Evaluation

TABLE OF CONTENTS

1

Introduction

Why the Push for Change

2

Major Findings

3

Different States, Different Approaches to...

Developing Teacher Evaluations

Using Teacher Evaluations

4

Bringing It All Together

5

Methodology

6

Acknowledgments

7

References



TOUCH OR
CLICK TO
NAVIGATE





1

Introduction

For decades, teacher evaluations were little more than a bureaucratic exercise that failed to recognize either excellence or mediocrity in teaching. Increasingly, this is no longer the case. Since 2009, the vast majority of states have made significant changes to how teachers are evaluated for the main purpose of improving instruction.

For most, this means evaluations will now take into consideration student achievement, as measured by standardized scores, alongside traditional methods like classroom observations, lesson plan reviews and others. Combined, these measures make for a more accurate assessment while providing valuable feedback to teachers on their strengths and weaknesses (Hull, 2011).

Although this paper will show that different states have taken different approaches to developing and implementing teacher performance systems, CPE discovered in a previous paper, ***Building a Better Evaluation System***, that there are certain key elements that should be part of all good teacher evaluations. Namely, broad stakeholder involvement throughout the process; the use of multiple measures, including data on student achievement and classroom practices; clear policies on how the information will be used; and the necessary resources and support to help teachers achieve those outcomes.





1

Introduction

Keep in mind that evaluating teachers, even partly on student achievement, is a relatively new concept and no system will be perfect. But by examining the different approaches states have taken, state and local education leaders can learn from each other to refine and improve their own systems.

Why the push for change?

Previous evaluation systems were inadequate. They identified nearly all teachers simply as satisfactory or not, without offering any useful feedback or direction to teachers on how they can improve.

New evaluation tools, notably improvements in classroom observations and the ability to connect teachers to student achievement, have been developed that are more accurate measures of teacher and school performance.

In response, federal programs such as Race to the Top and NCLB waivers provided incentives to states to redesign their evaluation systems.



**2****Major Findings**

- 1. Stakeholder input into the design of new evaluation systems has been important to gaining broad-based support.**
- 2. Most state evaluations include measures on how teachers impact their students' achievement.**
- 3. Improvements in classroom observations are a more accurate measure of instructional quality than previous classroom observation methods. They also provide valuable feedback to help teachers improve their own practice.**
- 4. Most states evaluate teachers on multiple measures to provide a more complete and accurate picture of a teacher's effectiveness. No state evaluates teachers on test scores alone.**
- 5. Most states are primarily focused on using evaluation for the purpose of raising teacher performance but also use the results to inform personnel decisions.**
- 6. Local school districts need flexibility in designing and implementing teacher evaluation systems so they are aligned to the needs of the district. But some also need strong support and resources from their states.**





3

Different states, Different approaches to...**Developing Teacher Evaluations**

Who's at the table?

Who calls the shots?

What goes into the evaluation?



**3**Different states, Different approaches to...**Developing Teacher Evaluations**

Who's at the table?

Inclusiveness is best

Including key stakeholders, especially teachers and administrators, in the design and implementation of evaluation systems is essential to creating an accurate measure of teacher performance and getting buy-in. School board members have a key official role as do elected state representatives and state department of education officials. Parents, business leaders, higher education officials and policy experts can also have an important place in this process.

Designing evaluation systems is complex and technical which is why 47 states left final decisions to experts based on feedback from constituents and officials. Only three states (Iowa, Michigan, & Utah) granted that authority to stakeholder panels. Nonetheless, it's vital to keep communication open and maintain trust during the process, especially when individual or group recommendations cannot be incorporated into the final plan.





3

Different states, Different approaches to...**Developing Teacher Evaluations**

Who calls the shots?

The balance between state vs. local determination varies by state

States vary by how involved they were in the design process:

- 1. High involvement:** 13 states mandated the requirements and components of the evaluation system and required districts to implement them with little flexibility.
- 2. Medium:** 17 states provided model evaluation systems that districts could either adopt or develop their own. For example, a state may mandate the use of student growth models and weights but still allow districts to decide the other features or choose alternative models as long as they meet the state criteria.
- 3. Low:** 21 states required each district to design their own system with state approval. The state may provide guidance but plays a small role in implementation.



**3**Different states, Different approaches to...**Developing Teacher Evaluations**

Who calls the shots?

In general, “medium involvement” seems to combine the best of the other two approaches. Districts are offered the option of a state model. This is especially helpful for districts that may not have the time and resources to design their own. Other districts are free to modify the state system or design their own to suit their unique needs.

However, one possible drawback to the lack of standardization within a state is that it can be difficult to compare teacher quality across districts.





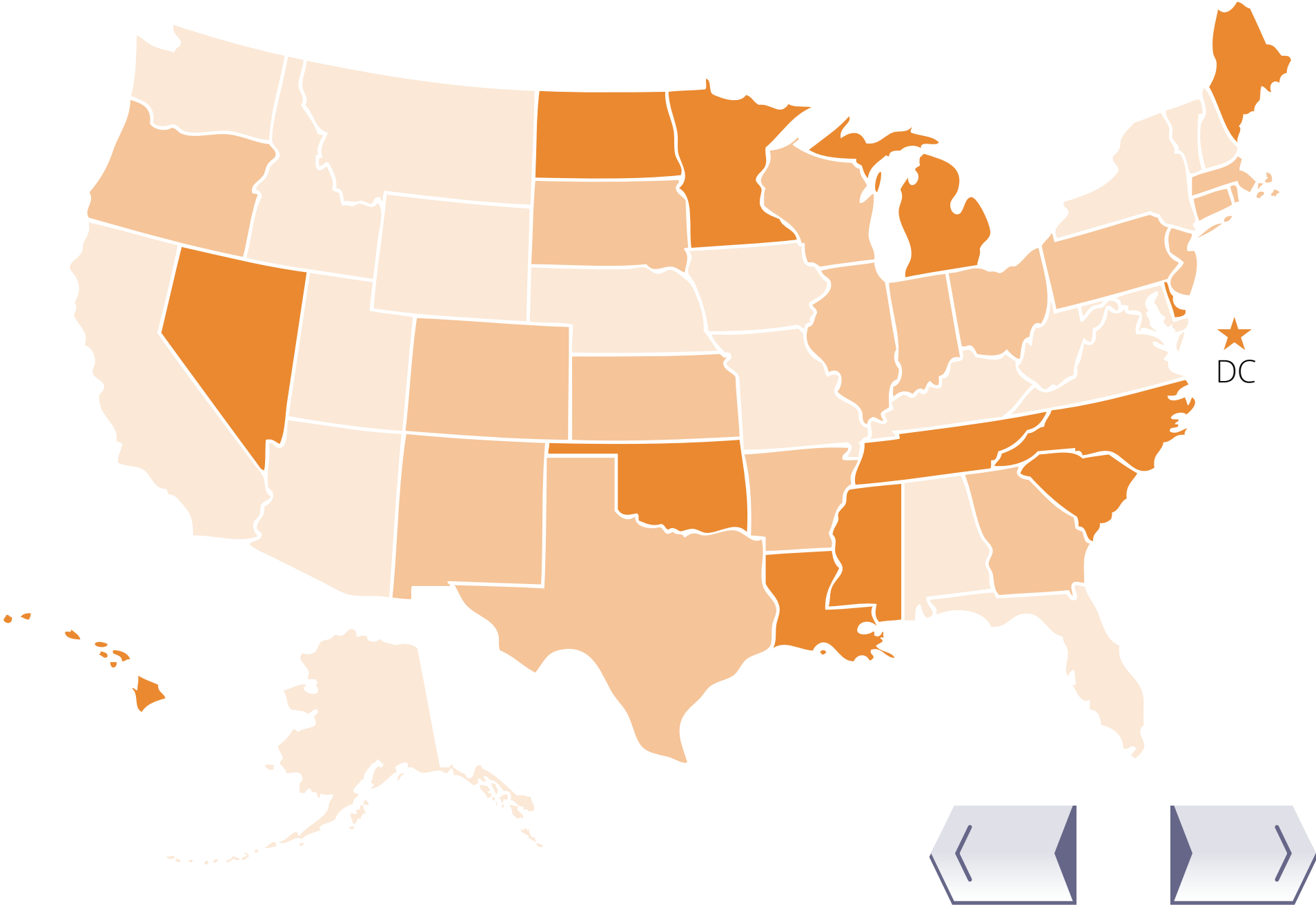
3

Different states, Different approaches to... **Developing Teacher Evaluations**

Who calls the shots?

Level of State Involvement

- high
- medium
- low



Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. State Roles



3

Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Most states use multiple measures

Forty-one states now require or recommend that teachers be evaluated using multiple measures of teacher performance. These include:

- **Student achievement data**
- **Classroom observations**
- **Other data: student surveys, lesson plan reviews, teacher self-assessments, and more**

States are now expected to evaluate teachers at least partially on the impact they have on their students' achievement. Value-added models (VAMs) are one of the best tools available for measuring teacher effect. However, VAMs are not perfect and don't provide teachers feedback on how to improve their performance. For this reason, all experts agree they should be used in concert with other measures of teacher quality.





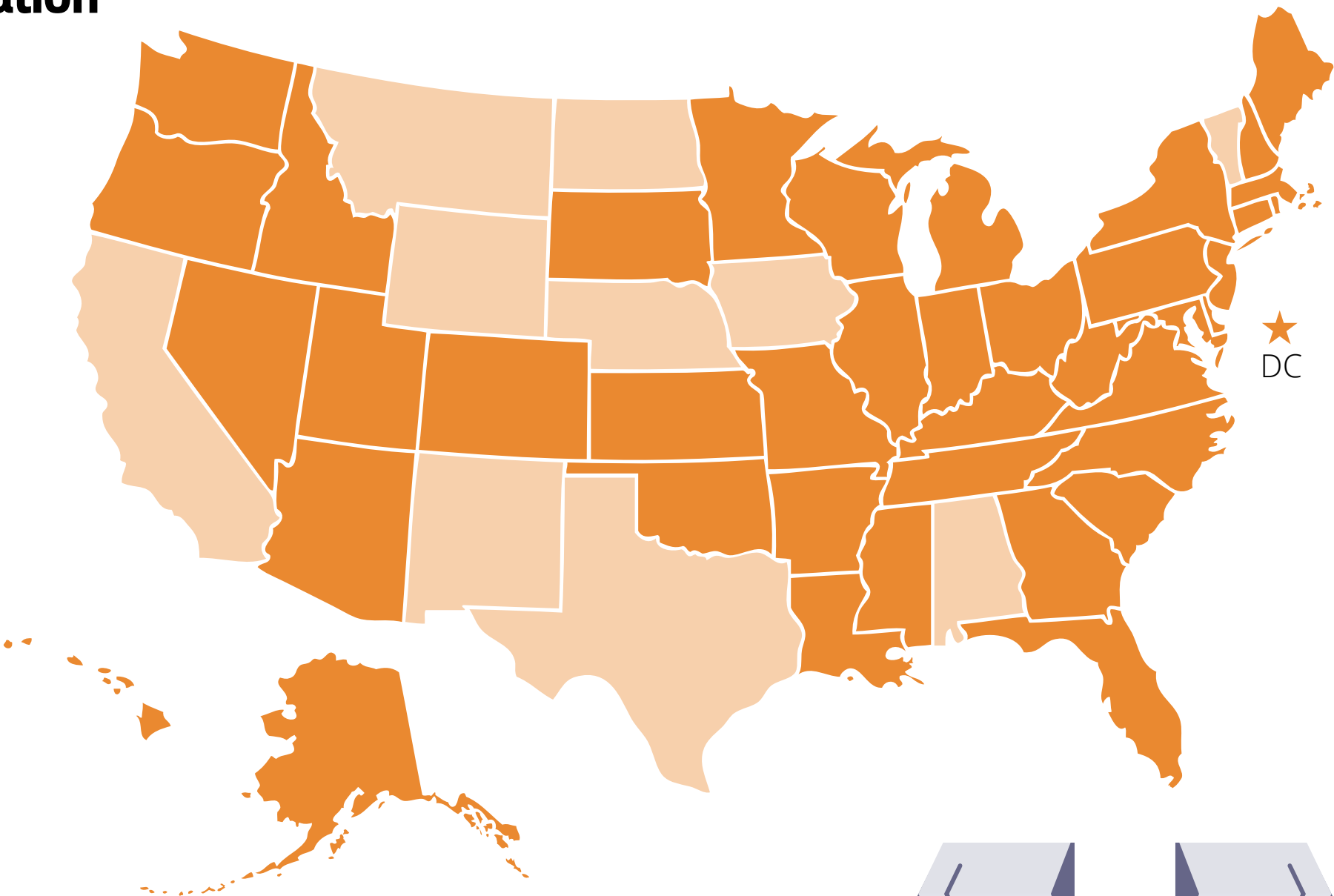
3

Different states, Different approaches to... **Developing Teacher Evaluations**

What goes into the evaluation?

41 states require or recommend using multiple measures in teacher evaluation

- Multiple measures
- Not specified



Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. Data Integrity





3

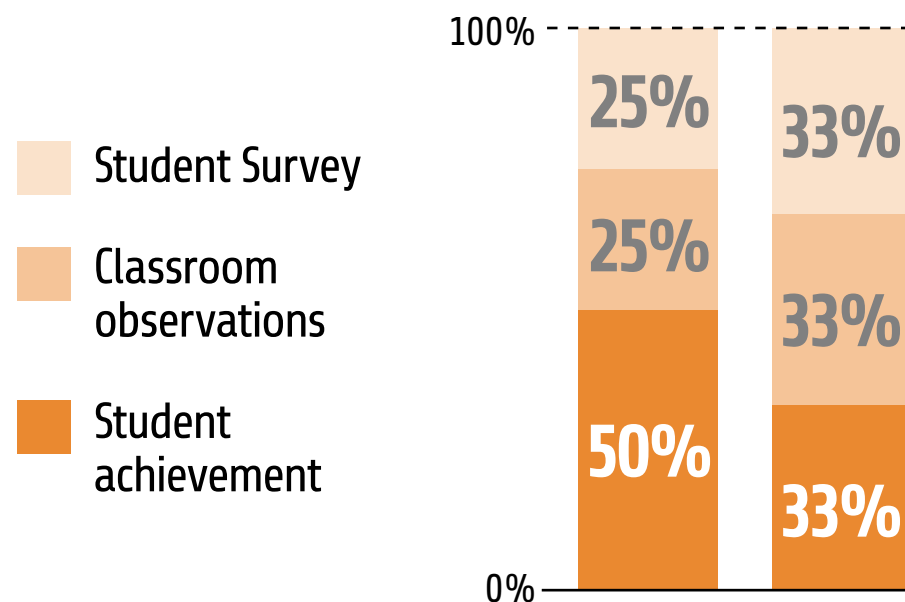
Different states, Different approaches to... **Developing Teacher Evaluations**

What goes into the evaluation?

No one way to combine measures

There is no magic formula for combining measures for evaluation. Research can provide some guidance but essentially it comes down to subjective decisions by state and/or district policymakers about what works best for their students and teachers. Such decisions should be informed, however, by experts and teachers.

Alternate ways to distribute the weight of different measures



The weight, or the numeric importance, given any measure should be driven by its quality (Baker, Oluwole and Green 2013). Formulas should also be field tested to determine if they represent a fair account of performance. If not, the formula should be modified.

The 2013 Measuring Effective Teaching study recommends two alternatives for distributing the weight of three factors as seen on the left.

The MET researchers argue that these formulas can provide an accurate measure of a teacher's performance without putting too much weight on any one measure which may incentivize teachers to focus too narrowly on that one aspect of teaching.

Source: Measures of Effective Teaching Project. Ensuring Fair and Reliable Measures of Effective Teaching. January 2013



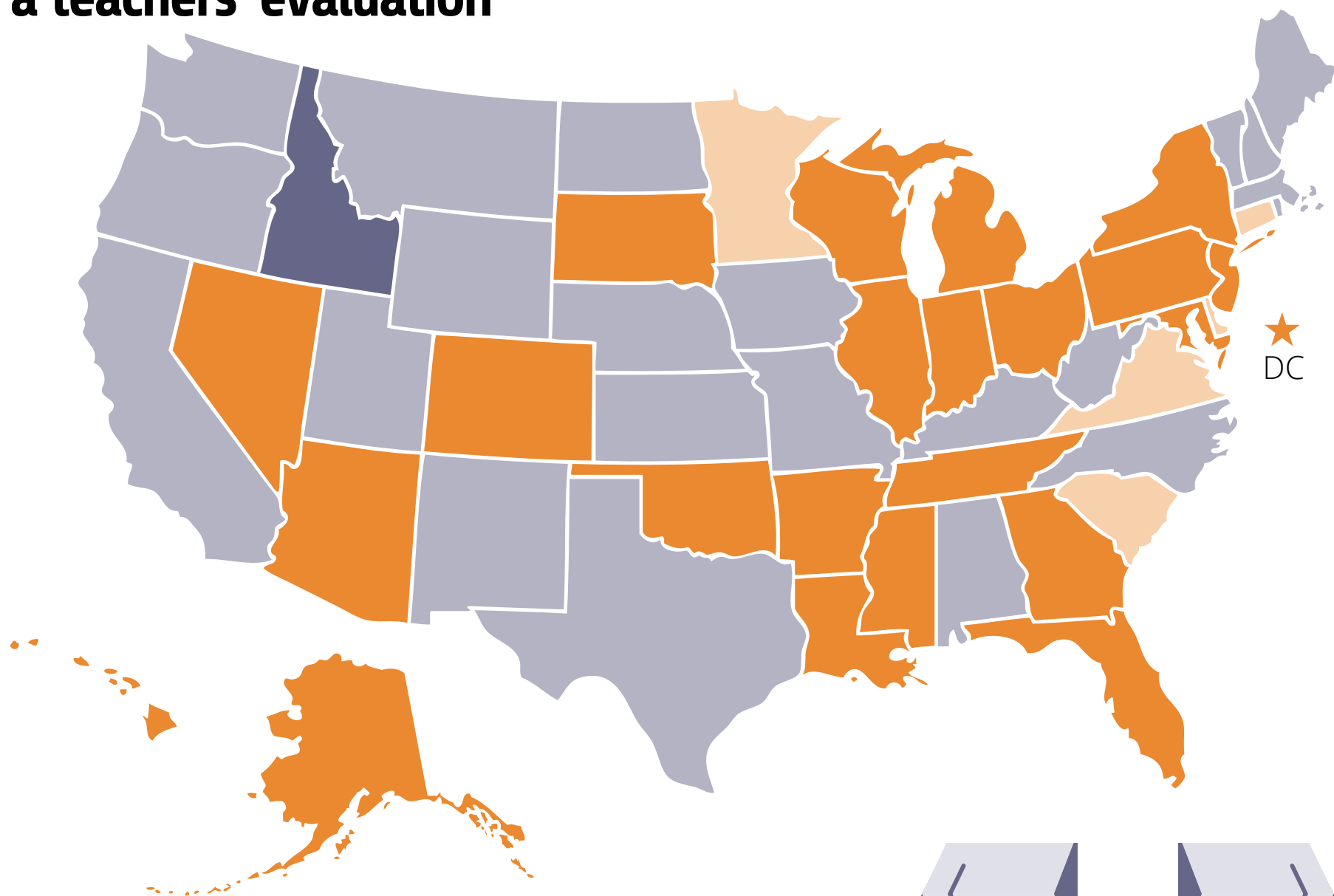
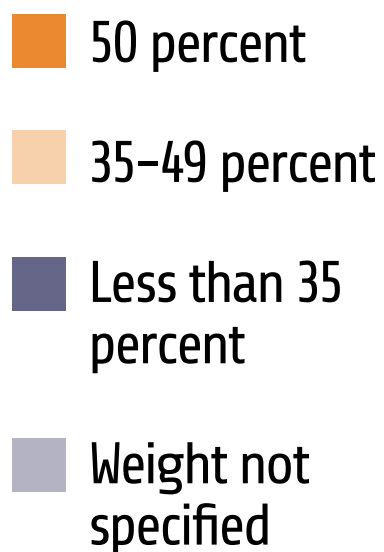


3

Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

23 states require or recommend that student achievement indicators comprise half of a teachers' evaluation



Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. Data Integrity



**3**Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Linking teacher performance to student achievement

Teachers' main objective is to grow student learning. Recognizing this, 38 states require districts to evaluate teachers, in part, on their impact on students' achievement. Another eight states recommend that they do so. Moreover, high quality student achievement measures have been shown to be objective and more accurate indicators of teacher performance than many traditional measures and are strongly correlated to other signs of a teacher's future effectiveness (MET 2013).

Even so, student achievement data does not capture the full impact teachers may have on their students. Likewise, measures like test scores do not represent a complete picture of what a student knows and understands -- yet another reason why multiple measures are important.





3

Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Standardized test scores are the most common, although not the only measure used to identify teacher effect on student learning. Statistical methods for linking scores to teacher performance can vary considerably but can be generally described in two ways, both of which attempt to capture student growth:

- **Value-added models (VAM):** Attempt to isolate the impact a teacher has on students' academic growth from other factors that impact student learning such as a student's socioeconomic status or their achievement on prior tests.
- **Student growth percentiles (SGP):** Measure how much progress a student has made relative to other students.





3

Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Value-added Model (VAM)

PROS

- Realizes not all students are likely to make the same growth from year to year
- Only model that attempts to isolate the impact a teacher has on student growth (Baker, Oluwole and Green 2013)
- More accurately identifies effective teachers compared to other measures including Student Growth Percentiles (Ehlert, et al. 2013)

CONS

- Typically utilizes complex statistical tools and multiple years of data, making them more difficult for non-statisticians to understand
- Research has shown VAMs aren't perfect measures of teacher effectiveness. Since the limitations are fairly well known there is resistance by a number of stakeholders to include them in any evaluation of teachers



**3**Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Student Growth Percentiles (SGP)

PROS

- Good measure of individual student growth from one year to the next
- Cheaper and easier to calculate plus easier to understand than VAMs
- More accurate at evaluating teachers than student test scores, which capture performance at one point in time
- Tend to be more popular with stakeholders than VAMs since their limitations are not as well known

CONS

- Does not attempt to isolate the impact a teacher has on student growth (Baker, Oluwole & Green 2013)
- Does not distinguish individual differences in rates of student growth
- Not as accurate at evaluating individual teachers as VAMs



**3**Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Linking teacher performance to student achievement

Most states also use other student achievement measures in evaluations as supplements to standardized test scores as well as provide data for teachers in untested subject areas and grade levels. These include Student Learning Objectives (SLO), formative assessments, district or teacher developed assessments, schoolwide value-added measures, and national assessments like AP, SAT, and ACT exams.

In most states districts have discretion over which alternative student achievement measures to include.

Student Learning Objectives are gaining in popularity, particularly for teachers who are not subject to standardized testing. SLOs are recommended parts of teacher evaluations in Connecticut, Indiana, Maryland, New Jersey, Nebraska and Rhode Island.





3

Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Student Learning Objectives (SLO)

PROS

- Incorporates teachers' professional judgment
- Useful for evaluating the progress of individual students
- Encourages teachers to reflect on the individual needs of their students
- Can be applied to all teachers, not just those in tested subjects

CONS

- They have not been used much on a large scale, so their effect in state or even district evaluation systems is not yet known
- Some concern that teachers may set lower goals for some students than others
- The SLO process is not standardized
- Cannot compare teacher results across districts or states





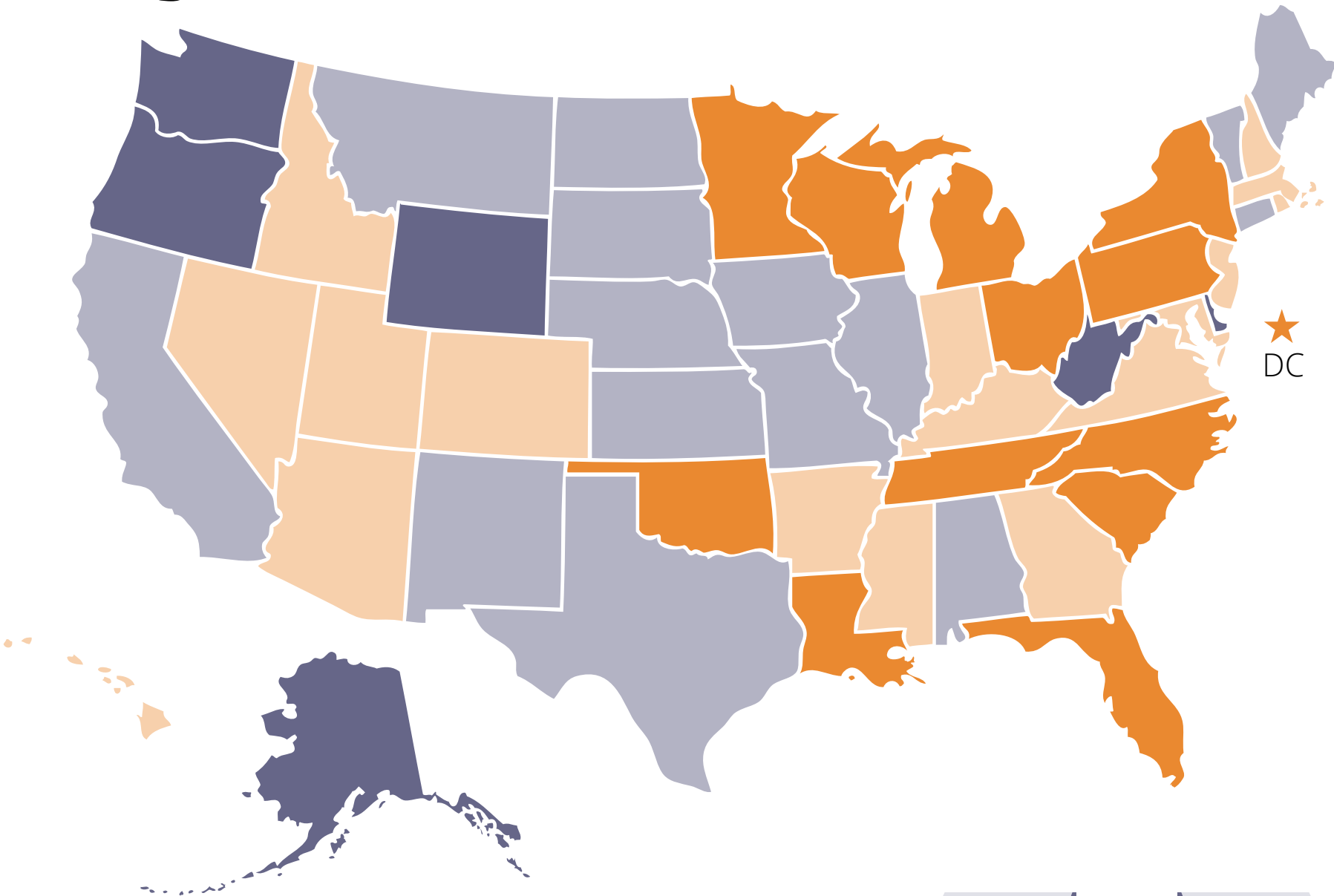
3

Different states, Different approaches to... **Developing Teacher Evaluations**

What goes into the evaluation?

Statistical models linking teachers and student achievement

- Value added
- Student Growth percentiles
- District choice
- Not specified/other



Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. Selecting Measures



**3**Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Linking teacher performance to classroom observations

Recent improvements in classroom observations provide a richer measure of the quality of a teacher's practice and give teachers valuable feedback about their performance.

The new research-based protocols identify practices most likely to affect student learning. Using well-trained observers further ensures that teachers being evaluated receive similar scores no matter who conducts the observations.

Research shows that observations are most effective when conducted multiple times per year.

In most states principals or their designee are responsible for conducting observations of all teachers. While staff evaluation is rightly a principal's responsibility, finding time to conduct a comprehensive observation of all teachers, even once year, can be difficult for these busy administrators.

Districts should consider training teams of observers. Indeed, other research found that observations are more accurate when they are conducted by more than one person.





3

Different states, Different approaches to...**Developing Teacher Evaluations**

What goes into the evaluation?

Other measures of teacher practice

Classroom observations are not the only measure of teacher practice. State and districts rely on other forms of evidence to get a fuller picture of a teacher's performance. These include:

- **student/parent surveys**
- **lesson plan reviews**
- **teacher self assessments**
- **measures of professional learning**
- **student artifacts**
- **teacher portfolios**

These other measures can provide rich, qualitative information about classroom instruction. As such they add to the body of feedback teachers receive on their individual strengths and weaknesses. However, not much is known about how accurate these other tools are at measuring the effectiveness of a teacher's practice.

The exception are well-designed student surveys, which recent research shows are strongly correlated to results on value-added measures and classroom observations (MET, 2013).

Whatever measures are used, districts should monitor their usefulness alongside other, more proven indicators to make sure that evaluations live up to their promise.





3

Different states, Different approaches to... **Using Teacher Evaluations**

Inform professional development

Identify low-performing teachers

Make personnel decisions

Report to the public

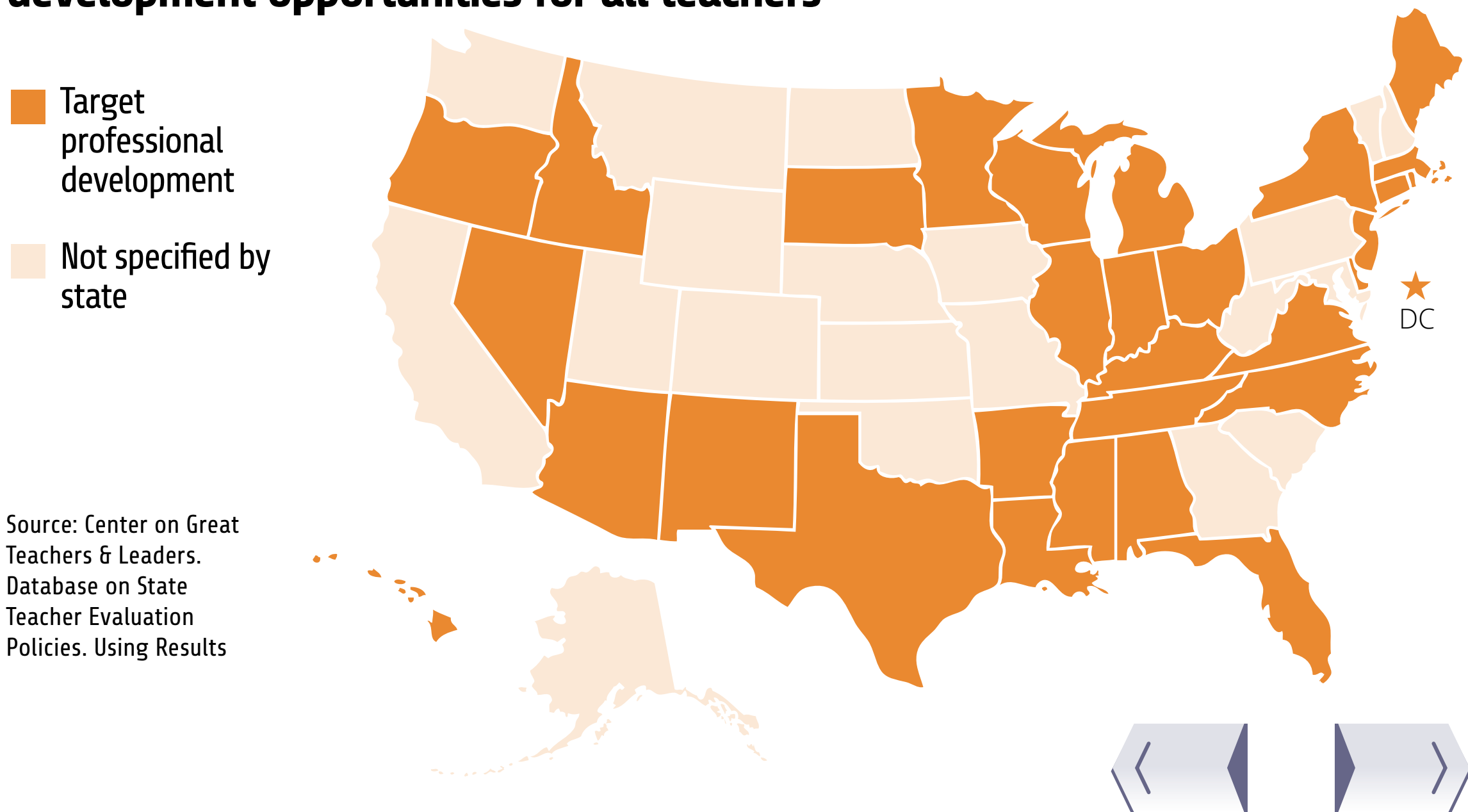


3

Different states, Different approaches to...**Using Teacher Evaluations**

Inform professional development

31 States align evaluation results to inform professional development opportunities for all teachers



Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. Using Results



3 Different states, Different approaches to... **Using Teacher Evaluations**

Identify low-performing teachers

Evaluation for continuous improvement

State Name	Interventions for low-performing teachers		
	Prof Dev/ Training	Improvement Plan	Mentoring/ Assistance
Alaska		X	
Arizona	X		
Arkansas		X	
Colorado		X	
Connecticut		X	
Delaware		X	
Florida	X		
Hawaii		X	
Idaho		X	
Illinois		X	
Indiana		X	
Iowa			X
Kentucky			X
Louisiana			X
Maine		X	
Maryland			X

State Name	Interventions for low-performing teachers		
	Prof Dev/ Training	Improvement Plan	Mentoring/ Assistance
Massachusetts		X	
Michigan		X	
Minnesota		X	
Nevada		X	
New Jersey	X		
New Mexico	X		
New York		X	
North Carolina		X	
Ohio		X	
Oklahoma		X	
Rhode Island		X	
South Carolina		X	
Tennessee	X		
Texas		X	
Utah		X	
Washington		X	

Few states require low-performing teachers to be dismissed

Just five states mandate dismissal for low-performing teachers. Even in these states, teachers are not dismissed until after they have received extensive support over multiple years.

In 27 states, teachers become eligible for dismissal after multiple low ratings. The decision to dismiss lies with the school district.

Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. Using Results





3

Different states, Different approaches to... **Using Teacher Evaluations**

Make personnel decisions

Most states also plan to use teacher evaluations for personnel decisions

Results from teacher evaluations can be a vital tool when making personnel decisions like tenure, promotions or dismissal.

Local discretion is essential when making such decisions, particularly when they have serious consequences. Such flexibility is important because even a well-designed comprehensive evaluation system cannot account for every factor that may impact a teacher's effectiveness. So, it's imperative that those closest to the school district make the final call.

- **28 states mandate the use of evaluation results in making personnel decisions.**
- **Another 12 states grant districts the discretion to use results for these purposes if they so choose.**





3

Different states, Different approaches to... **Using Teacher Evaluations**

Make personnel decisions

States allow the use of evaluations for personnel decisions

Teacher Compensation	Grant Tenure	License Renewal	Promoting Teachers	Teacher Assignments	Recognize Highly Effective Teachers
Colorado	Colorado	Delaware	Delaware	Arkansas	Colorado
Washington, DC	Delaware	Florida	Washington, DC	Delaware	Massachusetts
Florida	Hawaii	Hawaii	Illinois	Maine	Missouri
Indiana	Indiana	Illinois	Iowa	Rhode Island	New Hampshire
Maine	Kansas	Iowa	Massachusetts	Washington	
Massachusetts	Maryland	Louisiana	Michigan		
Michigan	Massachusetts	Maine	New Mexico		
Minnesota	Michigan	New Mexico	New York		
Nevada	New Jersey	North Carolina	North Carolina		
New Jersey	New York	Ohio	Oklahoma		
New York	North Carolina	Rhode Island	Oregon		
North Carolina	Rhode Island	South Carolina	Rhode Island		
Ohio	South Carolina		Tennessee		
Oklahoma	Tennessee				
Oregon					
Tennessee					
Utah					
Virginia					



**3**Different states, Different approaches to... **Using Teacher Evaluations**

Report to the public

Some states publicly report evaluation results

It's easy to argue that parents should have information about the quality of their child's teacher. It's harder to say that individual teacher evaluations should be public, however.

For one, a teacher's overall rating may not provide a full depiction of a teacher's true effectiveness. In addition, some experts assert that teachers are more likely to use their evaluation as a tool for continuous improvement if they know the results won't be publicized.

The public does have an interest in aggregate evaluation results at the school building level, as long as the identity of individual teachers is protected. It's important for the community to know how teacher quality is distributed throughout the district. Honest data is also useful for rallying public support for professional development and other teacher improvement efforts.

Note that exempting evaluation results for individual teachers from state open record laws may require amending the state constitution.





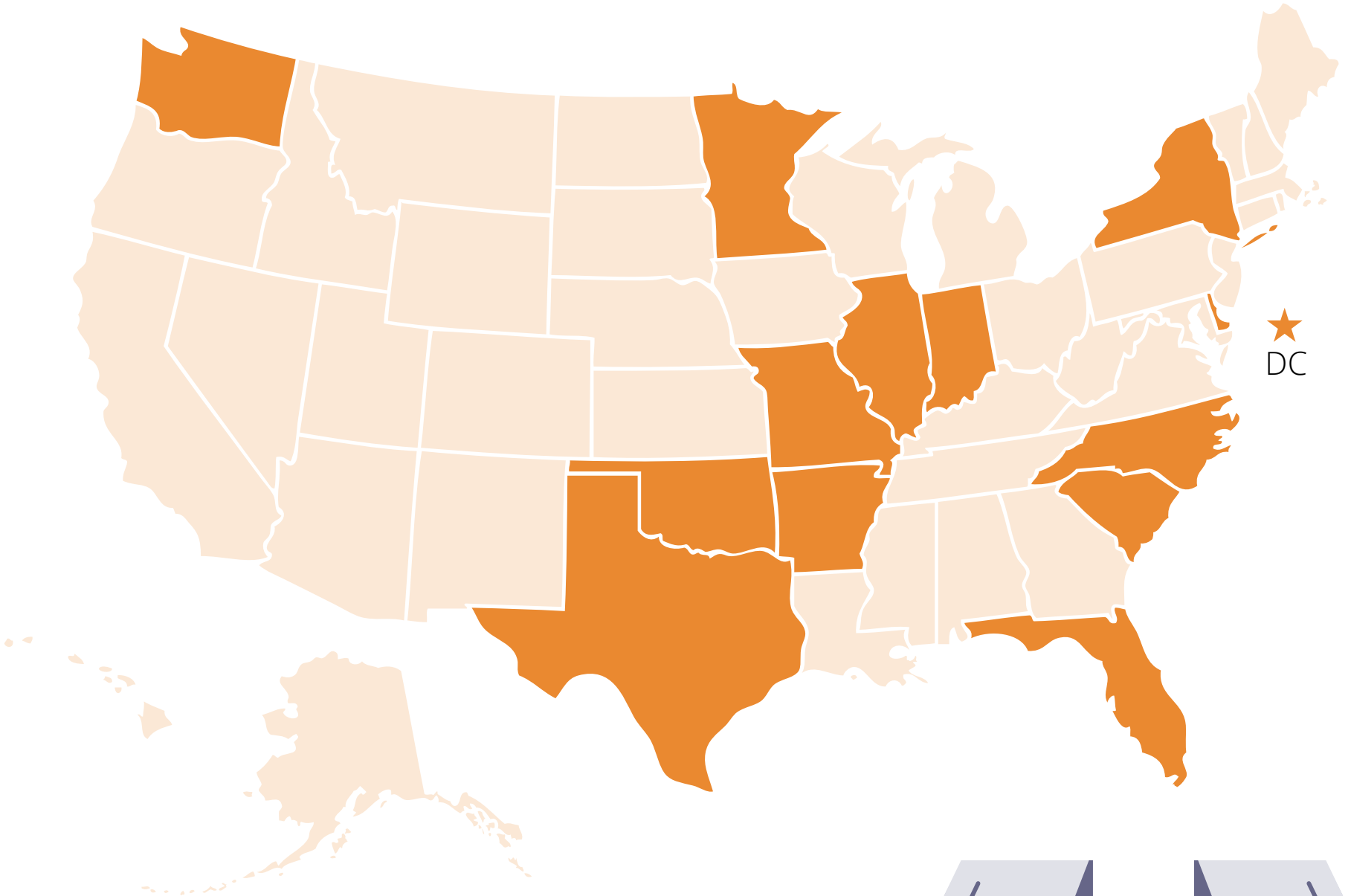
3

Different states, Different approaches to... **Using Teacher Evaluations**

Report to the public

14 states require aggregate evaluation data to be shared publicly

-  All public reporting
-  State does not require public reporting



Source: Center on Great Teachers & Leaders. Database on State Teacher Evaluation Policies. Data Integrity





3

Different states, Different approaches to... **Using Teacher Evaluations**

Report to the public

Public reporting does not typically include individual teacher data

Exempt from state open record laws

California · Delaware · Hawaii · Illinois · Indiana · Iowa · Kansas
Maryland · Massachusetts · Missouri · Pennsylvania · Rhode Island
South Dakota · Texas · Utah · Vermont · Wisconsin · Wyoming

Can release only with consent of the teacher

Alaska · Connecticut · Idaho · Louisiana

Can release with consent from third-party*

Arkansas · Kentucky · Mississippi · Montana · Nebraska · Ohio
Oklahoma · Oregon · Washington

Allow release under open records laws

Alabama · Arizona · Colorado · Florida · Georgia · Maine
Michigan · New Hampshire · New Jersey · New Mexico · New York
North Carolina · North Dakota · Nevada · South Carolina
Tennessee · Virginia · Washington, DC · West Virginia

*Such as record custodians, school district official, or state official.
SOURCE: Education Week. Access to Teacher Evaluations Divides Advocates. March 27, 2012.

**4**

Bringing it all together

Elements of a good teacher evaluation system

- Inclusive design and implementation process
- Policies on how information will be used
- Multiple measures of effectiveness including:
 - Data linking teachers and student achievement
 - Classroom observations
- Adequate resources and support





4

Bringing it all together

Districts need state support & local flexibility

States have a vital role in creating and implementing evaluation systems, but to be effective the systems should also be aligned to the unique needs of each district.

Most districts do not have the resources to design and implement a new evaluation system from scratch. Most will also need resources to support pilot testing, trained observers, professional development and other interventions to maintain a strong system.

Even so, evaluations will be most effective when they can at least be adapted to meet the unique needs of every district. This further allows districts to align evaluations to their goals.





4

Bringing it all together

Whether developing a teacher evaluation system, or implementing a new one, school district leaders should ask these questions:

How is the evaluation system developed?

What is the goal of the evaluation system?

Do those goals align with the district strategic plan?

What flexibility do districts have to tailor the evaluation system to the district's strategic plan?

Does the district have the knowledge and resources in-house to develop their own evaluation system or modify the state model?

Who was involved in development of the evaluation system? Were key stakeholders, particularly teachers, involved in some way?





4

Bringing it all together

What is included in the evaluation system?

What measures are included in the evaluation system? How much weight does each measure carry in the overall score?

How accurate are the results? What are the evaluation system's strengths and weaknesses?

What measures are used to determine the impact a teacher has on student achievement?

What statistical model is used to measure the impact? Why is that measure used? How accurately does it isolate a teacher's impact on student achievement?

Are the same measures used to evaluate all teachers? If not, how do they differ?

How often are teachers observed in a classroom setting? Does the frequency differ by experience or the teacher's previous performance level?

Do evaluators have enough time to conduct all the observations required without impeding on their other responsibilities?

Who conducts the observations? How are they trained?

Is the observation rubric researched-based and aligned with the district's goals?





4

Bringing it all together

How are results used?

When do teachers receive feedback from each observation?

Are the overall evaluation results used to improve instructional quality? If so, how?

Are the results used for personnel decisions? If so, how?

Are the results made public? If so, what information is made public?





Methodology

Data in this report is primarily based on information from The Center for Great Teachers & Leaders at the American Institutes of Research's (AIR) **Database on State Teacher Evaluation Policies (DSTEP)**. Data was collected in the summer of 2013. The database's collection methodology can be found **here**. In some cases additional resources were used to verify AIR's finding and to ensure the information was the most recently available. Due to differences in the timing of data collections as well as differences in the documents used the data in the DSTEP may differ from other reports of teacher evaluation systems.





6

Acknowledgments

- ▶ Jim Hull is the Senior Policy Analyst for the Center for Public Education. Research support provided by Allison Gulamhussein and Christine Duchouquette.
- ▶ The Center for Public Education is a national resource for credible and practical information about public education and its importance to the well-being of our nation. CPE provides up-to-date research, data, and analysis on current education issues and explores ways to improve student achievement and engage public support for public schools. CPE is an initiative of the National School Boards Association. **www.centerforpubliceducation.org**
- ▶ Founded in 1940, the National School Boards Association (NSBA) is a not-for-profit organization representing state associations of school boards and their more than 90,000 local school board members throughout the U.S. Working with and through our state associations, NSBA advocates for equity and excellence in public education through school board leadership. **www.nsba.org**

© October 2013 The Center for Public Education





7

References

- Baker, B. D., Oluwole, J. O. and Green, P. C. (2013). The Legal Consequences of Mandating High Stakes Decisions on Low-Quality Information: Teacher Evaluation in the Race-to-the-Top Era. *Education Policy Analysis Archives* 21(5).
- Ehlert, M., Koedel, C., Parsons, E. and Podgursky, M. (2013). Selecting Growth Models for School and Teacher Evaluations. National Center for Analysis of Longitudinal Data in Education Research Retrieved from <http://www.caldercenter.org/publications/calder-working-paper-80.cfm>
- Hull, J. (2011). Building a Better Evaluation System. Center for Public Education. Retrieved from <http://www.centerforpubliceducation.org/Main-Menu/Staffingstudents/Building-A-Better-Evaluation-System>
- MET. (2013). Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project's Three-Year Study. Retrieved from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf
- Sawchuck, S. (2012, March 27). Access to Teacher Evaluations Divide Advocates. *Education Week*.

