



SREB policy brief

Policy analysis and recommendations from the Southern Regional Education Board • November 2011

Measuring a Teacher's Value and Effectiveness in SREB States

In March, *National Journal* aptly pointed out, “It’s difficult to get more than 30 seconds into a conversation about education before teacher effectiveness comes up.” The current national fervor surrounding teacher effectiveness can be traced, in large part, to the priorities set by the federal Race to the Top (RTTT) grant competition in 2009 — and states’ responses to them. At a time when state budgets were shrinking significantly, RTTT offered states the chance to earn millions of federal dollars in exchange for their commitment to implement a new performance measure for teachers, among other key reforms. The grant competition clearly led to more than conversation; state policies changed. The National Council on Teaching Quality reported that 13 states nationwide — including six SREB states — had requirements by 2010 that student achievement results be the “preponderant criterion” — 50 percent or more — on teacher evaluations.

To win RTTT funding, each state was required to design an outcome-based measure that could differentiate among teachers using student achievement results to access their effectiveness in the classroom. This measure had to be a “significant factor” in overall teacher evaluations and used for “summative”

decisions (such as tenure and salaries). Yet, U.S. Secretary of Education Arne Duncan told members of the National Education Association in 2009 that “test scores alone should never drive evaluation, compensation or tenure decisions.” In doing so, he called for new models of teacher evaluation based on multiple measures.

In response to RTTT, seven SREB states — Arkansas, Florida, Louisiana, Maryland, Oklahoma, Virginia and Tennessee — passed legislation in either 2010 or 2011 that made an outcome-based performance measure a significant factor in teacher evaluations. In these seven states, the outcome-based performance measure counts (or will count) for 40 percent to 50 percent of overall teacher evaluations. Delaware accomplished essentially the same outcome through an administrative code change. (See Table 1).

A quick look back reveals how and why teacher effectiveness has become a federal and state priority and a major topic of public debate across the country. (For more discussion of recent teacher evaluation legislation in SREB states, see *Focus on Teacher Reform Legislation in SREB States: Evaluation Policies* at www.sreb.org.)

Setting the stage

Research in recent years has confirmed what many parents and principals have always known: Teachers are the most important factor in students’ education — and some teachers are better than others. This

realization is accompanied, however, by a tragic and now-substantiated reality: Students unlucky enough to be assigned ineffective teachers for three years in a row experience insurmountable academic losses.

SREB policy brief

Table 1

Recent Teacher Evaluation System Changes in SREB States

State	Citation	Year Passed	Year Implemented	Portion of Teacher Evaluation Based on Student Achievement ¹
Arkansas	House Bill 2178	2011	2014-2015	50 Percent
Delaware	Admin. Code Title 14 - 106A	2010	2011-2012	Significant Component ²
Florida	Senate Bill 736	2011	2011-2012	50 Percent
Louisiana	House Bill 1033	2010	2011-2012	50 Percent
Maryland	House Bill 1263	2010	2013-2014	Significant Component ³
Oklahoma	Senate Bill 2033	2010	2013-2014	50 Percent ⁴
Tennessee	Senate Bill 7005	2010	2011-2012	50 Percent ⁴
Virginia	House Bill 1500	2011	2012-2013	40 Percent ⁵

¹ The percentages listed are only for evaluations of teachers who teach subjects with annual state assessments.

² Requires teacher performance evaluations to include student growth data as a significant component weighted independently from the four other appraisal components, which generally comprise a group for rating purposes.

³ Requires teacher performance evaluations to include student growth data as a significant component. No component can count for more than 35 percent. See the SREB source, Page 6, for details.

⁴ The overall evaluation will consist of 50 percent student achievement growth data, made up of 35 percent student learning growth data and 15 percent other academic measures.

⁵ At least 40 percent of the evaluation will consist of student growth data.

Sources: *Focus on Teacher Reform Legislation in SREB States: Evaluation Policies*, SREB, 2011, and National Council on Teaching Quality, 2011.

In response, policy-makers have called for better teacher evaluations that can identify a range of teacher effectiveness — and researchers have responded with new measures and tools.

A 2009 report from The New Teacher Project, *The Widget Effect*, fueled this demand for better teacher evaluations. It found that fewer than 1 percent of teachers in the four states studied had received an “unsatisfactory” rating. The evaluations yielded almost no differentiation among the remaining 99 percent of teachers. So it was no surprise that almost every teacher received a “satisfactory” rating. In part, the evaluations failed to differentiate among teachers because principals — who conduct most teacher evaluations — lacked training in observing

and evaluating the quality of instruction, student learning or complex interactions between teachers and students. As a result, the evaluations were not useful in helping teachers improve their teaching.

Researchers have noted that teachers are evaluated too infrequently to yield useful results. Veteran teachers across the country generally are evaluated once each school year, and in some states even less often. New teachers, on the other hand, are evaluated from one to three times a year. To provide continuous feedback to teachers and to target professional learning, research suggests that evaluations be conducted at least three times a year for new *and* veteran teachers. (See Appendix A for state requirements on frequency of teacher evaluations).

The next generation of teacher evaluations needs to be comprehensive and able to help teachers improve their instruction, raise student achievement and provide school leaders with the information they need to evaluate them fairly.

In response to state requests for assistance in crafting policy on teacher effectiveness and evaluation, SREB has brought together key findings on the components of comprehensive teacher evaluation models from top researchers and policy groups.

Comprehensive evaluation models are needed

SREB states need to develop and implement comprehensive teacher evaluation models that use multiple measures. These evaluation models should provide teachers with information that can help them improve. The models also should differentiate effective teachers across a range of performance — rather than simply label them “satisfactory” or “unsatisfactory.”

Measures that can be included in a comprehensive teacher evaluation model include:

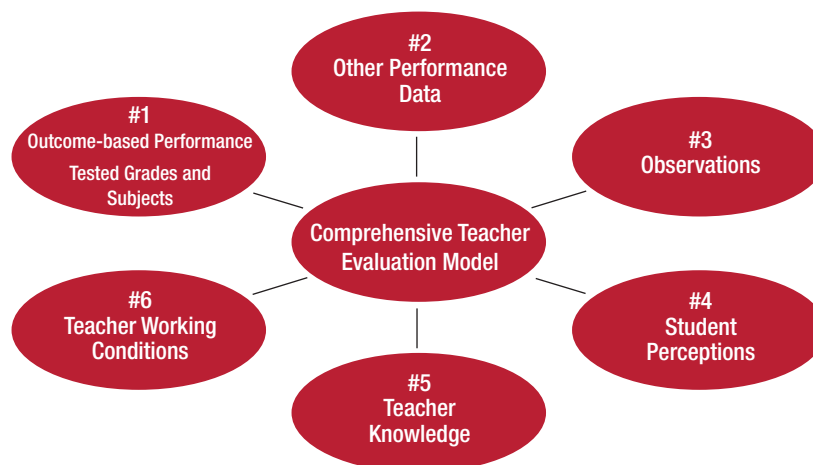
1. outcome-based student performance measures.
2. other student performance measures, as needed or required, particularly for teachers whose students are not assessed based on their students’ standardized test scores.

3. observations of teaching performance.
4. students’ perceptions of a teacher’s performance.
5. teacher knowledge — subject and pedagogical.
6. teachers’ perceptions of their working conditions, as needed.

Used collectively and appropriately, these measures can help states create a balanced comprehensive teacher evaluation system that will produce fair, reliable results. A strong, comprehensive model should incorporate both objective (outcome-based performance data) and subjective (observations and perceptions) measures, giving the overall model a better chance of predicting teacher effectiveness than if only one of these types of measures is used. (See Figure 1).

Figure 1

Components of a Comprehensive Teacher Evaluation Model Using Multiple Measures



SREB policy brief

Measure #1: Outcome-based performance measures

Outcome-based performance measures are based on complex statistics; they attempt to measure a teacher's impact on students' academic growth over time. Based on students' assessment results, the various models produce a positive or negative number that describes:

- a teacher's performance in relation to that of typical teachers.
- the average growth of students in typical teachers' classes.

Across all teacher performance ranges, outcome-based performance models are good at identifying the top 15 percent of most effective teachers and the bottom 15 percent of least effective teachers. They are weak at differentiating the vast majority of teachers who score in the middle — the close-to-average, typical teachers.

These models are used to evaluate teachers in core subjects because state assessment information is available in these subjects. In many states, they are used only with math and English/language arts teachers. Studies have shown they are also more accurate for math teachers than for reading and English/language arts teachers. (Students are more likely to learn math skills in school than reading and English/language arts skills. Learning in math is more easily attributable to individual teachers than is learning in reading and English/language arts.)

While these models provide the best single quantitative measure of teacher effectiveness, they come with some **limitations**.

- The models calculate effectiveness values for teachers using student assessment data each year. The students are *assigned* to the teachers each year. For all teachers to have fair odds of being assigned groups of students with a reasonable chance of making normal or "typical" progress in a year, class assignment needs to be made randomly to all teachers each year. For a wide variety of appropriate educational reasons, they are not. If summative decisions are attached to outcome-based performance, states will need to find ways to safeguard against problems that can arise if **student assignments are not fairly made**. States also need to assure teachers who are assigned atypical classes year after year (for example, math classes for English-language learners) that adjustments will be made in their overall evaluations.
- Class size and testing errors also can cause outcome-based results for an individual teacher to be unpredictable from year to year. The National Center for Education Statistics (NCES) reported that when using **only one year's worth of teacher-student data** to measure outcome-based performance results, there was a one in three chance that an average teacher could be misidentified as low-performing. To reduce teacher misidentification, NCES and other researchers recommend that outcome-based performance measures be based on three or more years of data. The chance of error is reduced to one in 10 with three years of data.
- Outcome-based measures in their purest, clinical sense are intended to attribute student growth to the teachers that fostered the growth. But tying student performance to a specific teacher or even to a group of teachers is difficult. Depending on the annual school calendar, middle grades and high school students typically have half a dozen teachers over the course of the year. Even in elementary schools, where students are generally assigned a "teacher of record," students also are served by a host of other teachers who work in collaborative teams. Yet teacher evaluation models are designed to produce evaluation values for the individual teacher alone. Researchers have yet to **find ways to untangle the growing "attribution knot"**. Schools and districts will need to struggle to answer the question: To whom can the school attribute student learning outcomes for students taught by teams of teachers?

Measure #2: Other performance measures

Other performance measures contribute to the overall evaluation by providing the data states need to evaluate teachers who teach in subjects and in grades for which standardized assessment data are not available. Only about 30 percent of the nation's teachers teach in grades and subjects in which students currently are assessed with standardized tests — most commonly, in reading and math. Currently, the states cannot use outcome-based performance measures to evaluate more than two-thirds of their teachers. As a consequence, states need to find other performance measures they can use to evaluate these teachers.

The U.S. Department of Education has provided federal funding to states winning RTTT grants for the development of state assessments for non-tested grades and untested subjects. All states are likely to

benefit from the experiences of these states. To measure teachers' contributions to student learning in untested grades and subjects, states should create measures that will be **comparable** from classroom to classroom, measure between **two points in time**, and are **rigorous**. Currently, several options are being considered:

- designing appropriate tasks to determine mastery in key subjects through the use of portfolios, products or projects.
- creating new tests for areas in which assessments do not exist.
- allowing teachers to select appropriate student learning objectives and to determine how to assess students' growth toward meeting those objectives.

Measure #3: Observations

Observations verify what is occurring in a classroom. They are the most frequently employed evaluation technique and still provide the most direct way to evaluate teacher-student interactions. When done well, observing a teacher can be a powerful tool. New research in the Cincinnati Public Schools by Harvard Professor Thomas Kane (who serves as deputy director for Research and Data, Bill & Melinda Gates Foundation) shows that observations conducted by *trained* observers — with observation protocols based on valid standards — can produce results that parallel findings from outcome-based performance values produced for the same teachers. Kane's Teacher Evaluation System is based on the work of researcher Charlotte Danielson. Good protocols are effective only if used by skilled evaluators, but evaluators require training and experience not available in most states. Currently, only a few states

across the country — for example, North Carolina — have strong evaluator training programs.

The absence of evaluator training is a substantial threat to the quality and accuracy of teacher evaluations based on teacher observations. The National Comprehensive Center for Teacher Quality reports that untrained evaluators can produce inaccurate and inconsistent results even when they use proven observation protocols. The bottom line: Evaluators need training that goes beyond the one-time training exercises reported in *The Widget Effect* as prevalent. The training also should ensure that evaluators clearly understand what the observation protocol is designed to measure and how to use it properly. Training that falls short of these marks will greatly increase the odds that the results of evaluations are flawed.

SREB policy brief

Measure #4: Students' perceptions of teacher performance

Students' perceptions of a teacher's performance contribute to comprehensive evaluation models in **three ways**: They confirm performance data results, inform the teacher's practices and target professional learning. When marketing experts want to understand what consumers think, they ask consumers. In education, students are the key consumer constituents, but until recently the tools did not exist to gather a student's accurate perceptions of a teacher's performance.

Using a survey instrument developed and refined by researchers at Harvard University and Cambridge Education over the past decade, the Measuring

Effective Teacher (MET) Project has gathered students' perceptions of teacher performance in seven areas that are critical to students' learning experiences: [How much does the teacher] care, control, clarify, challenge, captivate, confer and consolidate.

The MET Project's 2010 survey results point to a strong correlation between teachers' outcome-based performance data and their students' perceptions of their teaching abilities. Even student perceptions of one teacher's performance when he or she taught multiple courses in the same subject were consistent across various groups of students.

Measure #5: Teacher knowledge

Measuring teacher knowledge is an important component of evaluation models. It ensures that teachers are held responsible for possessing the necessary subject and pedagogical knowledge they need to teach effectively. Education historian Diane Ravitch aptly pointed out in 2003 that there was a "mismatch between teachers' academic preparation and the increasingly rigorous demands of the classroom." While this mismatch still exists, work continues at the national and state levels to ensure that teachers have a deeper understanding of the subject and pedagogical knowledge they need to be effective in the classroom.

Recognizing the need to address teacher preparation and knowledge, the Council of Chief State School Officers (CCSSO) revised its **core teaching standards** in 2009. Based on new teaching and learning research, the new standards require a deeper understanding of how to teach critical concepts and use

differing perspectives to engage learners in critical/creative thinking and collaborative problem solving. Over 40 states have adopted the core teaching standards.

To assess teacher's knowledge and ability, some states already have developed new teacher performance assessments. These measures require teachers to demonstrate their knowledge by documenting their instructional planning and teaching, videotaping and analyzing their lessons, and collecting and evaluating evidence of student learning. The results from these assessments are being used for initial teacher licensing recommendations in one SREB state — Kentucky — as well as in California, Colorado and Oregon. In Connecticut, these assessments serve as a gateway between initial and professional teacher licensure. These assessments hold promise for adding information to comprehensive models of teacher evaluation.

Measure #6: Teachers' perceptions of working conditions

Teachers' perceptions of their school working conditions should be an additional tool for states to include in a comprehensive evaluation model because they reflect school-level reasons for poor teacher effectiveness. Research demonstrates that school working conditions affect both teacher performance and student achievement. North Carolina Governor Beverly Perdue has said, "As a former teacher, I understand what an important part working conditions play in the success of the school, the teacher, and, ultimately, the student."

The gathering of teachers' anonymous perceptions of their schools began in North Carolina almost a decade ago with the North Carolina *Teacher Working Conditions Survey*. Since then, analysis of that data has revealed that there is a clear connection between teachers' perceptions of their schools and

the academic achievement of students on state assessments.

Ten states (including four SREB states) — Alabama, Colorado, Illinois, Kansas, Maine, Maryland, Massachusetts, North Carolina, Vermont and West Virginia — now administer surveys of teacher working conditions. Student results from these states parallel North Carolina's findings: Good school working conditions help recruit and retain effective teachers — and more effective teachers mean more successful students. Bill Ferrier, a regional Teacher of the Year in North Carolina, notes that "providing teachers with the time, tools and resources — including [the] quality leadership, professional development and flexibility that they need in order to be successful . . . with every child defines schools that have great working conditions."

On further reflection

In the end, comprehensive evaluation models are needed to provide direction for professional development and to provide clear and transparent indicators for administrators to use in making "summative decisions," such as pay increases, tenure and termination decisions.

Since these comprehensive evaluation models are new, states should consider pilot-testing them in several districts before implementing them statewide to gather feedback and to make adjustments where necessary. They should consider using evaluation findings to target professional learning and support for underperforming and average teachers **first**. By doing so, the evaluation process will support the ultimate goal of increased student achievement.

When comprehensive evaluation models are in place, economist Eric Hanushek recommends focusing teacher dismissal and replacement efforts on the bottom 5 percent to 7 percent of the teacher performance band. Replacing more than this proportion of teachers with better-performing teachers may not be a realistic goal because of the sheer number involved.

But **replacing even this small proportion of low-performing teachers** with average teachers could make a tremendous difference: The United States could close the achievement gap with Finland in the lifetime of a child born today — and increase U.S. annual growth by 1 percent of gross domestic product (GDP). This seemingly small increase could amount to \$112 trillion dollars in present value.

SREB policy brief

SREB states and school districts need to **develop strong communication plans** for their new teacher evaluation models. Explaining comprehensive teacher evaluation models will not be an easy task. Good communication will require making the concept understandable and meaningful to the teachers being evaluated, to the parents concerned about their children's education, and to the students being taught. It will require being proactive, getting out in front of the issues before the teacher evaluations are implemented and the data are released, and staying in front of the issues by engaging stakeholders. It will mean not relying on single events or one medium. Instead, it should utilize the bevy of media and events available: PTA and school board meetings, town halls, e-mail, websites, newspapers and social media. Former superintendent and author Robert J. Ramsey warns that “. . . when educators fail to communicate fully, misinformation, misinterpretations, misunderstandings and mixed messages can cause the system's wheels to spin or come off altogether.”

States and districts need to consider the legal implications involved, too. Since the vast majority of teachers will receive a comprehensive evaluation that does not contain an outcome-based performance measure — based on a statewide standardized measure — for the foreseeable future, states may need to **adopt two different approaches to teacher contracts:**

- one for teachers whose evaluations contain statewide standardized, outcome-based performance measures.
- another contract for those teachers whose evaluations do not include these measures.

While these new contracts will need to incorporate new teacher effectiveness measures, states need to ensure that other aspects of teachers employee responsibilities are included in evaluations and contracts. Positive outcome-based performance data already have been used successfully to defend teachers against dismissal cases based on non-academic reasons, such as tardiness and poor behavior.

Conclusion

There is no end in sight to the teacher effectiveness conversation. As RTTT winners begin implementing their new teacher evaluation models, the feedback from those states will certainly stimulate more conversation and possibly influence future state and federal

policy. Discussions at the national level would seem to suggest that future federal legislation is likely to continue to call for states to address teacher effectiveness directly. If so, the conversation should continue for some time.

Appendix A

State	State-Required Teacher Evaluations Each School Year	
	For New Teachers	For Veteran Teachers
Alabama	3 or More Times	No
Arkansas	3 or More Times	Yes
Delaware	Twice	No
Florida	Once	Yes
Georgia	Once	Yes
Kentucky	3 or More Times	No
Louisiana	Once	No
Maryland	Twice	No
Mississippi	Not Addressed	No
North Carolina	3 or More Times	No
Oklahoma	Twice	Yes
South Carolina	Twice	No
Tennessee	3 or More Times	No
Texas	Once	No
Virginia	Once	No
West Virginia	3 or More Times	No

Source: *Blueprint for Change: 2010 State Teacher Policy Yearbook*, National Council on Teaching Quality, 2010.

SREB policy brief

References

Buddin, R. *Measuring Teacher and School Effectiveness at Improving Student Achievement in Los Angeles Elementary Schools*. Study for *The Los Angeles Times*, 2011.

Ferrier, B. "Reflecting on Teacher Working Conditions." Teacher Leaders Network, Center for Teaching Quality, December 7, 2008.

Hanushek, Eric. "Valuing Teachers: How Much is a Good Teacher Worth?" *Education Next*, Summer 2011.

Kane, Thomas J. "Evaluating Teacher Effectiveness: Can classroom observations identify practices that raise achievement?" *Education Next*, Summer 2011.

Learning about Teaching: Initial Findings from the Measures of Effective Teaching Project. The Met Project, 2010.

Measuring Teachers' Contributions to Student Learning Growth for Nontested Grades and Subjects. National Comprehensive Center for Teaching Quality, March 2011.

Perdue, Governor Beverly Eaves. "North Carolina Announces Survey to Assess Teacher Working Conditions." *The Gov Monitor: Public Sector News and Information*, March 16, 2010.

Race to the Top Fund. U.S. Department of Education — <http://www2.ed.gov/programs/racetothetop/>.

Ramsey, J. Robert. *How to Say the Right Thing Every Time: Communicating Well with Students, Staff, and the Public*. Corwin Press, 2009.

Ravitch, Diane. "A Brief History of Teacher Professionalism." White House Conference on Preparing Tomorrow's Teachers, 2003.

"Secretary Duncan Challenges National Education Association to Accelerate School Reforms." U.S. Department of Education. Press Release, July 2, 2009. Accessed at — <http://www.ed.gov/news/press-releases/secretary-duncan-challenges-national-education-association-accelerate-school-re-0>.

Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies. National Council on Teaching Quality, 2011. Accessed at — <http://www.nctq.org>.

The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness. The New Teacher Project, 2009.

Southern Regional Education Board

592 10th Street, N.W.

Atlanta, Georgia 30318

(404) 875-9211

www.sreb.org