

RAND RESEARCH AREAS
 CHILDREN AND FAMILIES
 EDUCATION AND THE ARTS
 ENERGY AND ENVIRONMENT
 HEALTH AND HEALTH CARE
 INFRASTRUCTURE AND TRANSPORTATION
 INTERNATIONAL AFFAIRS
 LAW AND BUSINESS
 NATIONAL SECURITY
 POPULATION AND AGING
 PUBLIC SAFETY
 SCIENCE AND TECHNOLOGY
 TERRORISM AND HOMELAND SECURITY

Using Student Performance to Evaluate Teachers

Over the past 15 years, a growing body of research has demonstrated that teachers are the most important school-based determinant of student achievement. Taking advantage of improved educational data systems that link students to their teachers, this research has used a class of statistical techniques called value-added models to estimate teachers' impact on their students' standardized test performance. Such research has also shown that many current teacher evaluation systems, which rely mainly on supervisor judgments, do not adequately reflect variation in teachers' estimated ability to raise test scores. As a result, there is a growing interest in developing teacher evaluation systems that incorporate value-added estimates of teachers' effects on student achievement.¹

However, systems that incorporate student test scores into teacher evaluations face at least two important challenges. First, they must support *valid* and *reliable* inferences about teachers' contributions to student learning, and, second, they must attempt to include teachers who do not teach subjects or grades that are tested annually. The Center for American Progress, with support from the Bill and Melinda Gates Foundation, asked RAND to review the literature and examine how five educational systems are beginning to approach these challenges.

Choosing Which Student Performance Measures to Use

Policymakers should take technical quality considerations into account when using student achievement data to inform teacher evaluations. One such consideration is *validity*, which refers to the accuracy and appropriateness of the inferences about student achievement drawn from a test score. Another is the *reliability* of the test scores, or the extent to which scores are consistent

Key findings:

- Teacher evaluation systems should incorporate multiple measures of teacher performance.
- Student assessments used in these systems must support valid inferences about teacher effectiveness.
- For teachers of nontested subjects or grades, evaluations may include supplemental student assessments or school- or department-level student performance.
- Estimates of teachers' effectiveness become more stable as they incorporate additional years of student assessment data and are averaged over multiple years.
- New teacher evaluation systems should be monitored to minimize unintended consequences.

over repeated measurements and free of measurement error. Threats to the validity of value-added estimates include an undue focus on test preparation in lieu of better teaching of the underlying content, as well as differences in the content tested from one year to the next. Validation studies that compare student performance on high-stakes and low-stakes assessments and that examine the grade-to-grade alignment of content on standardized tests can gauge the severity of these threats.

It is also important to consider the reliability of teachers' value-added estimates obtained from student test scores. Studies conducted by RAND and others have shown considerable year-to-year instability in teachers' value-added estimates, which is at least partially due to measurement error. Such error can be reduced by including multiple years of student test scores in the estimation models and averaging teachers' value-added estimates across multiple years.

¹ Value-added estimates use a statistical technique that combines student test scores over multiple years to estimate the impact of individual schools or teachers.

This product is part of the RAND Corporation research brief series. RAND research briefs present policy-oriented summaries of published, peer-reviewed documents.

Corporate Headquarters
 1776 Main Street
 P.O. Box 2138
 Santa Monica, California
 90407-2138
 TEL 310.393.0411
 FAX 310.393.4818

© RAND 2011

Measuring Performance in Grades and Subjects That Are Not Tested Annually

Many states test students in reading and math only, and only in grades 3 through 8, because those are the requirements of the federal No Child Left Behind Act. These states can therefore estimate the value-added effectiveness of only a subset of their teachers. To address this limitation, some educational systems are purchasing commercial assessments or developing local assessments to use in nontested grades and subjects, though it is important that the tests be developed and validated for this purpose. Some systems use aggregate student performance measures, such as schoolwide averages on reading and math tests. However, doing so can mean that some teachers are held accountable for the collective performance of a grade, department, or school, while others are held accountable primarily for the performance of the students they teach. Also, aggregate student performance in core subjects may not reflect the contributions of teachers of specialized subjects such as physical education or world languages.

Additional challenges arise when students lack prior test scores or are enrolled in a teacher's class for only part of the school year. It is prudent to estimate teachers' value-added effects on achievement using only the students who were in their classes most or all of the year and who have prior test scores on record.

Lessons for Developing Performance-Based Teacher Evaluation Systems

RAND researchers used publicly available documents to profile two states and three districts that have begun or are planning to incorporate measures of student performance into their teacher evaluations: Denver, Colorado; Hillsborough County, Florida; Washington, D.C.; and the states of Tennessee and Delaware. Based on the literature and insights gleaned from those systems, the study generated five recommendations for policymakers who are working to develop new teacher evaluation systems.

Create comprehensive evaluation systems that incorporate multiple measures of teacher effectiveness. Because any single measure of teacher effectiveness is prone to measurement error, and because not all teachers teach annually tested grades and subjects, it is important to evaluate teachers across multiple dimensions. Evaluations should be based not only on value-added estimates that use student test scores but

also on other data sources, such as observations of teachers' classroom practices and evidence of their contributions to their schools.

Attend not only to the technical properties of student assessments but also to how the assessments are being used in high-stakes contexts. Because teacher evaluations can have far-reaching consequences, it is important to use measures with strong technical properties. Particularly when supplementing the state accountability tests, it is important to make sure that the additional tests remain reliable and valid when used for the high-stakes purpose of evaluating teachers' performance.

Promote consistency in whatever student performance measures teachers are allowed to choose. In some cases, educational systems supplement state accountability tests by allowing teachers of nontested subjects and grades to choose the student assessments for which they will be held accountable. In these cases, it appears helpful if teachers are given clear parameters about the measures they can choose and if the supplemental measures chosen are consistent across classrooms.

Use multiple years of student achievement data in annual value-added estimation and, where possible, average teachers' annual value-added estimates across multiple years. Research shows that using multiple years of student achievement data when estimating teachers' value-added effectiveness increases the accuracy and precision of those estimates. In addition, averaging teachers' value-added estimates across multiple years reduces the instability of the estimates.

Find ways to hold teachers accountable for students who are not included in their value-added estimates. Students who have spent only part of the year in a teacher's classroom or who lack prior test scores cannot easily be included in their teachers' value-added estimates. Yet, teachers can be encouraged to demonstrate their impact on these students' achievement in ways that do not depend on test scores, such as by demonstrating improvements in students' performance on classroom assignments.

Educators May Benefit from Being Flexible and Willing to Learn

Incorporating student performance measures into teacher evaluation systems is a complex undertaking. As they strive to enhance teacher evaluation systems, policymakers may benefit from examining what other systems are doing and learning from their struggles and successes. ■

This research brief describes work done for RAND Education documented in *Incorporating Student Performance Measures into Teacher Evaluation Systems*, by Jennifer L. Steele, Laura S. Hamilton, and Brian M. Stecher, TR-917-CAP, 2010, 51 pp. (available at http://www.rand.org/pubs/technical_reports/TR917.html). This research brief was written by Sydne Newberry. The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.

RAND Offices

Santa Monica, CA • Washington, DC • Pittsburgh, PA • New Orleans, LA/Jackson, MS • Boston, MA • Doha, QA • Abu Dhabi, AE • Cambridge, UK • Brussels, BE



EDUCATION

CHILDREN AND FAMILIES
EDUCATION AND THE ARTS
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INFRASTRUCTURE AND
TRANSPORTATION
INTERNATIONAL AFFAIRS
LAW AND BUSINESS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
TERRORISM AND
HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Support RAND

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Education](#)

View [document details](#)

Research Brief

This product is part of the RAND Corporation research brief series. RAND research briefs present policy-oriented summaries of individual published, peer-reviewed documents or of a body of published work.

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).