



School of Education, University of Colorado at Boulder
Boulder, CO 80309-0249
Telephone: 802-383-0058

NEPC@colorado.edu
<http://nepc.colorado.edu>

GETTING TEACHER ASSESSMENT RIGHT: WHAT POLICYMAKERS CAN LEARN FROM RESEARCH

Patricia H. Hinchey, Pennsylvania State University

Everyone agrees that teacher quality makes a big difference in student learning. The practical puzzle facing policymakers is, *what's the best way to determine who's a high-quality teacher?*

A review of 275 articles and reports on the subject of teacher assessment found conclusively that placing excessive emphasis on test scores alone can undermine the goal of developing an excellent teaching force.

The current enthusiasm for using student test scores as the sole measure of teacher effectiveness stems from several sources, including convenience. Test score data are readily available because of No Child Left Behind Act requirements, and non-statisticians often perceive statistical analyses as objective, simple and reliable.

Today, “value-added” assessments of teachers using such scores are becoming more widespread, despite a steady stream of warnings from prominent researchers who caution that attempts to directly attribute students’ scores to their teachers are inherently flawed. These flaws cause teachers to be judged excellent one year and awful the

What is Value-Added Modeling?

Interest in assessing teachers based on student test scores has intensified with the development of value-added modeling, which increases the capacity of researchers to isolate the effect of a single teacher from other influences on student achievement, such as prior teachers, home influences, school environment, and student motivation. This modeling, sometimes known as value-added assessment, uses complex formulas to estimate students’ likely achievement gains in a given year. Actual gains are compared to this estimate, and classroom teachers are credited (or blamed) when students show greater (or lesser) gains than expected. While representing a technical advance, value-added modeling has clear weaknesses that limit its capacity to evaluate individual teachers.

next. Because of this and other significant problems that arise when very high stakes are attached to test scores, experts have repeatedly warned against the primary use of value-added modeling to make high-stakes decisions.

Other assessment methods, such as teacher observations, portfolios, and self-reports on classroom practice, can help create the needed balance of information. Because every tool for assessing teacher performance has both strengths and weaknesses, and because assessment can have multiple goals, it is better to develop a comprehensive assessment system than to adopt a single, dominant measure of performance. Several such balanced systems have already been developed, and experience suggests they have promise for helping to nurture and promote a highly skilled teaching staff.

Based on the research reviewed, the brief that this summary draws from recommends that policymakers employ a full-spectrum assessment system that continually improves teaching but also enables the timely dismissal of teachers who cannot or will not improve. Steps that policymakers can take toward that goal include:

- **Involve all key stakeholders, especially teachers and administrators, in system design.**
- **Rather than employing a single assessment tool, gather evidence from multiple sources. Combine strategies so that the weakness of any single tool is offset by the strengths of another.**
- **Be sure that the criteria for assessing performance, artifacts, or other factors are credible and are well understood by teachers and assessors.**
- **Be clear about the purposes of any assessment before selecting strategies. Where formative and summative assessments are to be combined, plan to address the challenges of dual-purpose systems.**
- **Provide high-quality, ongoing training for assessors, and routinely calibrate their efforts to ensure consistent application of criteria.**
- **Look to high-quality research on existing tools and programs to inform the design of assessment systems.**
- **Commit sufficient resources to produce high-quality, productive assessment.**

To be clear, the evidence shows that teacher performance can be assessed. Like any other assessment, however, the more limited the inputs, the more limited the applicability of the results.

For the full 30-page research brief, including citations, please visit:

<http://nepc.colorado.edu/publication/getting-teacher-assessment-right>

*The mission of the **National Education Policy Center** is to produce and disseminate high-quality, peer-reviewed research to inform education policy discussions. We are guided by the belief that the democratic governance of public education is strengthened when policies are based on sound evidence. For more information on NEPC, please visit <http://nepc.colorado.edu/>. The policy brief was made possible in part by the generous support of the Great Lakes Center for Education Research and Practice (greatlakescenter.org).*