

Using Student Tests to Measure Teacher Quality

Teaching matters. Study after study confirms that students who have high quality teachers make significant and lasting achievement gains. Those with less effective teachers play a constant game of academic catch-up. Having high quality teachers consistently tops parents' list of critical elements of an effective school. As the National Commission on Teaching and America's Future stated, "What teachers know and can do makes the crucial difference in what students learn." But what's a "good" teacher? And how do we measure effective teaching?

It turns out that assessing teaching quality is a very complex task. Research suggests that good teachers know the subjects they teach and are skilled at communicating these subjects to students. But even this description is rather vague. Isn't there a simple way to judge whether or not a teacher is effective?

Some have suggested that there is such a method—use students' test scores. After all, their argument goes, tests are supposed to measure student learning. So don't test scores measure how well teachers teach? They point out that large-scale tests are an increasingly important part of the education landscape and that the federal *No Child Left Behind Act* requires that all students in grades 3 – 8 be tested annually in reading and mathematics. So, they suggest, why not use test scores as the means to gauge teacher effectiveness? This assessment brief takes a closer look at this issue.

Some Challenges in Using Test Scores to Measure Teacher Quality

Testing experts continue to debate the wisdom—and fairness—of evaluating teachers on the basis of their students' test scores. While some advocate and some oppose this approach, there is reasonable consensus about some of the challenges presented by using test scores as the *sole* measure of teacher quality. These challenges include the following:

- **Test scores show patterns over time.** While 2-3 or more years of test scores might reliably show growth or lost educational ground, a single year of scores – which often is suggested as the yardstick, is not a reliable measure of teacher effects.

Resources

Helpful Web Sites

CAESL<http://www.caesl.org/>**CRESST**<http://www.cse.ucla.edu/>**WestEd**<http://www.wested.org/>**Parent Portal at LHS**<http://lhsparent.org>**Greatschools.net**<http://www.greatschools.net/>**National PTA**<http://www.pta.org/>**National Parent Information Network**<http://www.NPIN.org/>**Family Education Network**<http://www.familyeducation.com>



- **Tests provide only a sample of student performance.** Large-scale tests typically show only what students can do on a paper and pencil exam, but don't measure other important types of student learning, such as writing and problem-solving ability.
- **Tests don't necessarily measure what teachers are required to teach.** Surprising as it may seem, tests do not always measure what teachers are supposed to be teaching. For example, many large-scale tests measure basic skills, but teachers may be required to teach to a defined set of academic standards.
- **Context matters.** Many factors influence student achievement and, therefore, affect student test scores. These include home support, school attendance, family income level, and parents' level of education. The result is that teachers in wealthier communities are likely to "look" better because their students are likely to score higher on tests. But this is often more a measure of students' home environments than of teachers' instructional effectiveness.

So what's the solution? Clearly, test scores are important. They can't and shouldn't be ignored. It's simply not reasonable to suggest that test scores have no place in evaluating teacher effectiveness.

A Promising Approach

The real question is, "What do teachers contribute to student learning and how do we know?" In other words, is there a way to use student test scores to determine how much of student learning can be attributed to a particular teacher? In fact, there is a method that has been proposed. Pioneered in Tennessee, it's called "value-added." It's called "value added" because it's meant to be a measure of the value teachers add to students' learning. Using this method requires some complicated analysis of statistics, but the end result is considered to be useful information – based on test scores – about teachers' contribution to student learning. Also in this system, each student is compared to his or her own performance the previous year, so other issues (such as family income) aren't involved in the analysis. Tennessee's results from this system are noteworthy. These results include:

- **Teachers make a difference.** While students may start at comparable achievement levels, they can end up in quite different places as a result of the sequence of teachers to which they are assigned.
- **Teacher effects are cumulative.** After three years of not very effective teachers, students may never be able to catch up academically.
- **The most effective teachers produce the greatest learning gains in all students.** The most effective teachers produced the greatest learning gains in the *lowest* achieving students, followed by average achievers, and then by above average students.

Some Cautions to Bear in Mind

In designing a way to measure teacher effectiveness, it is important to keep a couple of things in mind. First, the goal of evaluation is *improvement*. Yes, good teacher evaluation systems should weed out the small number of “rotten apples.” But the main goal of evaluation ought to be pointing out areas where teachers need improvement and then providing support so they can become more effective. Second, even when using a system such as “value-added,” test scores should be **just one component of a comprehensive evaluation system**. Such a system ought also to include observations of actual classroom teaching and examination of student work.

What You Can Do

- Given that tests are being used for a greater number of purposes than ever before, it's important that parents keep themselves informed. Because most student achievement tests were not designed to be used for teacher evaluations, we recommend they be reviewed and validated for such purposes before such use.
- As with any evaluation or assessment that has important consequences for students, teachers, or schools, we encourage the use of multiple methods—more than a single test score—when seeking to evaluate teacher quality.
- Find out how your child's school and district evaluate teacher quality. Are there continuing opportunities for teachers to grow professionally? Do these professional development programs have evidence to demonstrate that they have a positive impact on teacher growth and student achievement?
- Become better informed through reading some of the resources listed below.

Useful Resources

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Baker, E.L., Linn, R.L., Herman, J.L., and Koretz, D. (Winter 2002). *Standards For Educational Accountability Systems*. (CRESST Policy Brief 5). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Herbert, L., Perry, M, Liebig, G., & Teaque, J. (April, 2002). *Teacher Pay in California: Is it fair? Is it competitive? Is it enough?* Palo Alto, CA: EdSource, Inc.

Koppich, J. “Spotlighting Teacher Quality: A Review of Teacher Evaluation,” (1998) Paper prepared for the K-12 Program Policy Council of the American Federation of Teachers.

Koppich, J. “Investing in Teaching” (2001). Washington, D.C.: National Alliance of Business.



Lewis, A. (2001). 2000 CRESST Conference Proceedings: *Educational accountability in the 21st century*. (CSE Technical Report 549). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing. pages 11-15.

Rosenthal, L. *What Makes a Great Teacher?* GreatSchools.com: <http://www.greatschools.net/cgi-bin/showarticle/ca/259/parent->

Julia E. Koppich the main author of this brief, is an evaluation specialist who was kind enough to contribute her expertise on this topic to the Public Understanding strand of CAESL. CAESL reviewers included: Jacquye Barber, Ron Dietel, Joan Herman, and Lincoln Bergman.

Note: This article was developed by the Public Understanding strand of CAESL to summarize basic information for parents and the general public. It is not a CAESL position statement nor does it necessarily represent the precise views of diverse reviewers. We welcome comments!

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This material is based upon work supported by the National Science Foundation under Grant No. 0119790.

© CAESL 2003. All rights reserved. Permission to reproduce, with CAESL copyright notice included, is hereby granted.