

## Test Failure

Andrea Orr [[http://www.epi.org/authors/bio/orr\\_andrea/](http://www.epi.org/authors/bio/orr_andrea/)]

August 27, 2010

For public school students, back-to-school means not just long hours in a classroom, but a tough regimen of tests designed to measure their proficiency in basic subjects like math and reading.

Increasingly, these same tests are also being used to evaluate teachers. In 2001, Congress approved the No Child Left Behind Act, requiring all schools to measure student performance through standardized tests. The volumes of data that have been collected since then have also been used to “grade” teachers, determine pay raises, and sometimes terminate them. More recently, schools have applied a method of value-added modeling, or VAM, in order to apply a more sophisticated analysis to the test score results.

But even when value-added modeling is used in the analysis, student test scores are not reliable indicators of teacher effectiveness, according to the EPI report, *Problems with the Use of Student Test Scores to Evaluate Teachers*. The paper was co-authored by a group of distinguished education scholars and policy makers, including four former presidents of the American Educational Research Association, a former assistant U.S. Secretary of Education, EPI Research Associate **Richard Rothstein**, and others. The authors find that the accuracy of these analyses of student test scores is highly problematic. They argue that the practice of holding teachers accountable for their student’s test score results should be reconsidered.

The paper’s publication comes at a time of increased use of test score results to measure teacher performance. The Los Angeles Unified School District recently applied value-added modeling to test results to evaluate teachers and published the results in *The Los Angeles Times*. In Washington D.C., School Chancellor Michelle Rhee has said she would consider making these value-added assessments public as well. And U.S. Secretary of Education Arne Duncan has called for all school districts to make such results public.

“If new laws or policies specifically require that teachers be fired if their students’ test scores do not rise by a certain amount, then more teachers might well be terminated than is now the case,” the authors state. “But there is not strong evidence to indicate either that the departing teachers would actually be the weakest teachers, or that the departing teachers would be replaced by more effective ones.”

The EPI paper finds that student test scores, even with value-added modeling, cannot fully account for a wide range of factors such as students’ background and the “learning loss” that often occurs over the summer. In fact, while students overall lose an average of about one month in reading achievement over the summer, lower-income students lose significantly more. The value-added modeling also cannot take into account the influence of student’s other teachers, including previous teachers and teachers of other subjects, as well as tutors.

The authors also stress that an excessive focus on the basic math and reading skills that are the focus of standardized tests can lead to a narrowing of school curriculums, at the expense of subjects such as science, history, the arts, civics, foreign languages, writing, and research.

And, they dispel the notion that private-sector employees have long been subject to a similar sort of quantitative performance review.

“Rather, private-sector managers almost always evaluate their professional and lower-management employees based on qualitative reviews by supervisors; quantitative indicators are used sparingly and in tandem with other evidence,”

they state. "Management experts warn against significant use of quantitative measures for making salary or bonus decisions."