

EDUCATION WEEK

Published Online: April 6, 2010

Published in Print: April 7, 2010, as **Tying Teacher Evaluation to Student Achievement**

COMMENTARY

Tying Teacher Evaluation to Student Achievement

Caution, Yellow Light Ahead

By Susan H. Fuhrman

The Obama administration, through its Race to the Top initiative, is encouraging states to develop approaches for evaluating teachers that incorporate student-achievement results. This aspect of the program has been controversial, prompting some teachers' unions to refuse to endorse state applications for competitive federal grants. However, a number of efforts to develop such indices of teacher effectiveness are under way, and the American Federation of Teachers' president, Randi Weingarten, has publicly endorsed including student-achievement results along with other measures to evaluate teacher success.

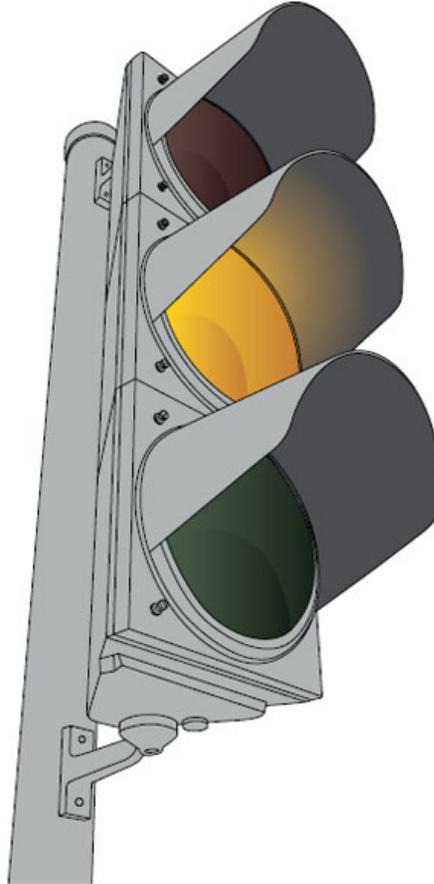
It is likely, then, that some form of teacher evaluation linked to student achievement will play a significant role in a number of upcoming policy initiatives. It is therefore critical, in order to ensure fairness to teachers, that any plans to reward or punish them for gains their students have or have not made control for differences among students in their family situations and other factors that are beyond the teachers' control. The best method for ensuring that evaluation includes such controls is called the value-added approach.

Recently, the [National Research Council](#) and the [National Academy of Education](#) jointly issued [a report](#) on value-added approaches, based on findings from a November 2008 workshop funded by the Carnegie Corporation of New York and co-sponsored by the NRC and the NAE. The report's goal was to provide policymakers with an improved understanding of the potential role of value-added methodologies, given their known strengths and weaknesses, so that officials could then better decide whether (and how) to implement them in their jurisdictions.

According to the report, "value-added models" refer to a variety of sophisticated statistical techniques that measure student growth and use one or more years of prior student test scores, as well as other background data, to adjust for pre-existing differences among students when calculating contributions to student test performance.

Current accountability systems rely predominantly on the "percent of children reaching proficiency," which educational measurement experts call a "status" measure. Schools making good progress but not yet reaching desired average levels of achievement are not rewarded, and schools with high-achieving students have no further incentive to improve if they've already reached the mandated proficiency level. Workshop participants were generally positive about adding measures of growth to status measures in accountability systems.

They voiced less support, however, for using value-added measures for high-stakes decisions, especially about individual teachers. One reason is that it is currently impossible to use test-score gains for the large number of teachers whose students are not given standardized tests, including those teaching in the earliest grades and those in subjects like art, music, and social studies, where standardized tests are not routinely used. And it would



—iStockphoto.com/blackred

[Back to Story](#)

Let them play it out
before they live it out

TRY IT FOR FREE! ENTER CODE: [edweek](#)



**WILL
CAMPUS**

Fully Interactive Games that will Improve
Your Teens' Decision Making Skills

"Policymakers need to move carefully in adopting any approach—value-added or otherwise"

be most unfortunate if attempts to improve teacher accountability exacerbated one of the most criticized aspects of current accountability systems, namely the overreliance on standardized tests.

**added or otherwise
—in making
important decisions
about individual
teachers."**

Moreover, even in the cases where tests already exist, such as for teachers of reading and mathematics in grades 3-8, value-added approaches raise significant concerns. Recent research suggests that they give an accurate picture of teacher-related gains in achievement only if students are randomly assigned to teachers. But if, for example, administrators systematically assign struggling students to the "best" teachers (as may be the case in many schools) or to new, inexperienced teachers (as is the case in many other schools), those teachers' measured gains relative to those of their colleagues will likely suffer.

There are a number of other concerns about the implementation of value-added models, including the following:

- Many tests cover sufficiently different content from one grade to the next that score gains do not have the same meaning across grades. Many state assessments, in fact, are not scaled to measure grade-to-grade growth or to make growth comparisons.
- Value-added estimates for a teacher can fluctuate for a variety of reasons, many not necessarily related to actual effectiveness at producing student gains on achievement tests. For example, high turnover of students throughout the year can affect the gains students make on achievement tests; and, if the class size is small, the scores of only a few students can affect the size of the gains. These kinds of errors can be reduced—but not eliminated—if administrators take several years of teacher performance into account when making important decisions.
- Factors other than an individual teacher's efforts affect student performance in any given year. These include the efforts of other teachers involved with a student, the extent of support the student receives outside of school in completing homework and learning the material (tutoring, parental help, and the like), and other family and societal factors that might influence student achievement.

The lesson of the NRC-NAEd report is that even though value-added methodologies offer a number of advantages over other approaches that consider test-score data in a vacuum, policymakers need to move carefully in adopting any approach—value-added or otherwise—in making important decisions about individual teachers. Value-added approaches hold great promise, but there is a need to develop better tests (and other thoughtful measures of student learning) and better measures of teacher practice to use along with test scores, so they are not the sole factor used to evaluate teacher effectiveness.

Susan H. Fuhrman is the president of the National Academy of Education (NAEd) and also the president of Teachers College, Columbia University.

Vol. 29, Issue 28, Pages 32-33

RELATED STORIES

["Standards Aren't Enough,"](#) October 8, 2009.

["Universities and Public Schools: A Partnership Whose Time Has Come,"](#) November 12, 2008.

["If We're Talking About Race, Let's Talk About Education,"](#) May 7, 2008.