

THE VALUE OF VALUE-ADDED MEASURES

Summary Findings:

- Value-added measures represent an attempt to separate the contribution of the teacher to student learning from all of those other things (such as the prior learning of the student) that the student already brings to the table.
- Statistical value-added models can generate predicted year-end scores for each student based on individual characteristics, and then rate teachers based on whether their students over-perform or under-perform.
- Opponents of value-added measures are concerned the use of value-added measures will lead to higher levels of public scrutiny of schools and educators.

The concept of value-added measures of teacher or school effectiveness is prompting a great deal of discussion in K-12 Education policy circles. This debate reached a boiling point last year when the Los Angeles Times published a database of the value-added scores for all teachers in the nation's second largest school district. Proponents argue value-added measures provide important information on school and teacher effectiveness. Opponents argue value-added measures are imprecise instruments which measure student background instead of teacher or school quality. The purpose of this policy brief is to provide the reader with a general understanding of the concept of a value-added measure as well as the potential benefits and perils of more widespread use of such value-added measures.

WHY DO WE NEED VALUE-ADDED MEASURES?

Simply put, the concept of value-added attempts to measure the value that each individual teacher (or school) adds to the learning of his or her students during a given time period. This is an important distinction from basing educator effectiveness solely on year-end student performance. Many have criticized the idea of rating a teacher based on where his or her students ended up, because this takes no account of where the students started. Thus, evaluating teachers only on how well students performed at year-end could lead to an over-rating of teachers who are assigned very capable students or an under-rating of teachers assigned a struggling group of students.

Indeed, our reliance on bad measures of school effectiveness illustrates the need for value-added indicators. The well-known *Adequate Yearly Progress (AYP)* school rating system is an overly-simplistic measure that can generate inappropriate conclusions. Under AYP, schools are rated based on the fraction of kids meeting a particular performance level. It is easy to see that schools full of already-advanced students may well meet AYP without adding much educational value at all. Conversely, a school serving struggling students might add a great deal of educational value by helping the students progress a great deal throughout each school year. Nonetheless, this effective school still might not meet the pre-set AYP bar. Thus, measures like the AYP that do not focus on growth, or "value-added", can lead to multiple problems. Based on AYP, we might overlook very effective teachers and heap unwarranted praise upon less effective teachers, who were simply given the already high-achieving students.

Value-added measures represent a genuine attempt to create more meaningful indicators by separating the contribution of the teacher from all of those other things (such as the prior learning of the student) that may well influence student performance.



**UNIVERSITY OF
ARKANSAS**

THE OFFICE FOR EDUCATION POLICY
211 GRADUATE EDUCATION BUILDING
FAYETTEVILLE, AR 72701
(479) 575-3773 (OFFICE)
(479) 575-3196 (FAX)

OEP@UARK.EDU

[HTTP://OFFICEFOREDPOLICY.COM](http://OFFICEFOREDPOLICY.COM)

OEP DIRECTOR:

Gary W. Ritter, PhD

RESEARCH ASSOCIATES:

Nathan C. Jensen

Caleb P. Rose

GRADUATE FELLOWS:

James L. Woodworth

Jeffery R. Dean

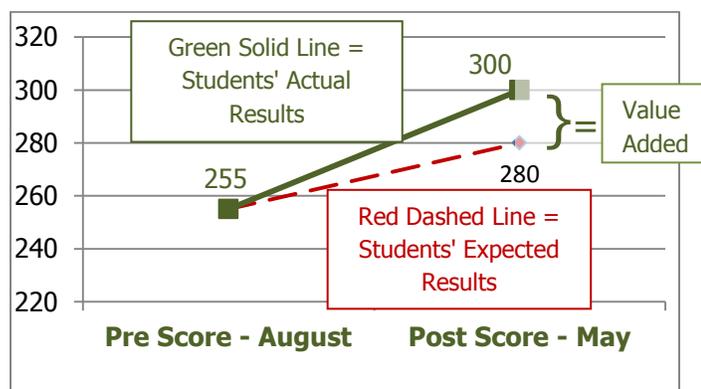
TYPES OF VALUE-ADDED MEASURES

There are two broad strategies for measuring the value-added for a teacher.

Simple Method. The most straightforward strategy to measure the value-added by teachers is a simple pre-post growth measure based on student performance on standardized assessments. In the example below, the teacher would earn a value-added score of +45 because her students grew by an average of 45 points during the school year.

POST-Score		PRE-Score		<i>Teacher Value-Added Score</i>
Average Score of Students in May	-	Average Score of Students in August	=	
300	-	255	=	+ 45

Sophisticated Method. More complicated statistical models have been developed to predict student performance on standardized assessments based on prior performance and on background characteristics believed to influence student achievement, such as gender, race, or socioeconomic status. Using such predictive models, researchers can examine whether students do better or worse than they would be expected to do. Teachers whose students exceed expectations earn high value-added scores. In the example below, the teacher would earn a value-added score of +20 because the teacher's students earned an average score of 300 rather than their "expected" score of 280.



IS SIMPLER ALWAYS BETTER?

To be sure, student year-end results alone are poor measures of teaching effectiveness. Beyond that, however, there are both strengths and weaknesses associated with more simplicity. Simple pre-post growth measures work best when the tests are given twice during a school year -- at the beginning and at the end. If these measures are based on year to year testing, they may ignore learning that occurs (or does not occur) during the summer months. Furthermore, pre-post value-added measures based on year to year assessments require vertical equating. That is, the test scores must be scaled across grades so that they are comparable from year to year.

While more sophisticated statistical models offset some of these problems, there are other problems that arise. Foremost among these is the lack of transparency -- evaluation systems work best if all stakeholders have a true understanding of the process. Thus, if statistical models are used in teacher evaluations, it is critical that school leaders and teachers are fully able to understand the evaluation process. Translating statistics into understandable language is not an easy task.

CRITICISM OF VALUE-ADDED MEASURES

Not all education observers have embraced value-added measures with open arms. In August 2010, several prestigious researchers published a briefing paper for the Economic Policy Institute (EPI) articulating their concerns with the use of value-added measures for teacher evaluation. Here are a few of the criticisms

1. Value-added models are too imprecise and unstable to be used as a basis for high-stakes decisions.
2. Teacher evaluations based on value-added models may discourage teachers from working with neediest students.
3. Ratings based on value-added test scores may encourage school leaders to focus the curriculum too heavily on tested subjects.
4. Publication of value-added ratings may demoralize teachers and discourage collaboration.

Given the weaknesses in current rating systems for schools along with the criticisms directed at value-added ratings, what should policymakers do?

DO VALUE-ADDED MEASURES HAVE A PLACE IN TEACHER ASSESSMENT?

A few months after the publication of EPI research brief, a set of education researchers at the Brown Center on Education Policy at Brookings published what was, essentially a rebuttal. This group, also composed of esteemed academics, begins by acknowledging that value-added measures are indeed estimates of teacher effectiveness and thus contain a level of measurement error. They conclude, however, that the measures need not be perfect to be useful. Ignoring the information provided in value-added models, they argue, does no good.

In our view, this argument is more convincing than that of the critics. Instead of focusing on the fact that value-added measures are not perfect, the discussion should revolve around whether or not value-added measures are better than the teacher evaluation strategies currently in use. Indeed, we believe that value-added measures have the potential to be much better than existing strategies used for teacher evaluation.

CONCLUSION

At this point in time in most schools across the country, teacher effectiveness is measured primarily by principal evaluations based on informal interactions and a few classroom observations. According to a recent report, *The Widget Effect*, a study which includes data for some school districts from Arkansas and other states, "94 percent of teachers receive one of the top two ratings and less than 1 percent are rated unsatisfactory." While it would be nice to believe all teachers are far above average, it is difficult to imagine any occupation where less than 1 percent of employees are unsatisfactory. Clearly, current methods of teacher evaluation are not generating the information that school leaders need. To guide our school systems effectively, administrators must assess which teachers are enhancing student learning for all kids.

Value-Added Models: Our View

- While value-added models are not perfect, they are likely more useful than are current methods of teacher evaluation.
- Value-added models that incorporate multiple years of data and make reasonable allowances for measurement error can be attentive to the genuine concerns of critics about the instability of these measures.
- Value-added models that account for where students start can certainly recognize the positive contributions made by teachers and schools which serve needy populations.
- If done well, value-added models need not discourage collaboration. Indeed, models that incorporate school-wide measures of student achievement can encourage collaboration.
- It is possible that value-added models may demoralize some teachers; it is also possible that the existing regime in which highly effective teachers go unrecognized demoralizes teachers.
- The level of simplicity of value-added models should vary, depending on the availability of data, the testing schedule, and the relationship between teachers and leaders. In any event, the measures need be as clearly communicated as possible to all stakeholders.
- While student test scores are not the only measures of student learning, they are objective indicators and are based on curricular standards. Good value-added models will certainly rely heavily on these scores.
- Test based value-added models, combined with other measures such as improved principal observations, would likely represent a substantial improvement over the current practices of teacher evaluation.

For more information about this policy brief, please contact the Office for Education Policy at oepe@uark.edu or visit us on the OEP Blog at <http://officeforedpolicy.com>