

EDUCATION WEEK

Published Online: November 16, 2004

Published in Print: November 17, 2004

'Value Added' Models Gain in Popularity

Growth Yardstick Appeals to States

By [Lynn Olson](#)

The concept sounds appealing: Measure the effectiveness of schools and teachers based on the amount of academic progress their students make from one year to the next. Often known as "value added" measures because they track the "value" that schools add to individual students' learning over time, such methods are increasingly popular with educators and policymakers.

 [Back to Story](#)

Some view the methods as an antidote to accountability systems that focus solely on getting children to a specified achievement level on a state test, regardless of where they start. Others view them as a way to isolate the effects of teachers and schools on learning, separate from such background characteristics as race and poverty.

Three national conferences on the topic took place last month alone. And this week, the Washington-based Council of Chief State School Officers planned to host a meeting on their use.

"Value-added measurement is a very active area today," Nancy S. Grasmick, the state superintendent of education in Maryland, said during a conference at the University of Maryland College Park last month. "We know there's controversy surrounding this," she added. "We need to ferret out all of the factors and not just jump into this without a strong research base."

Indeed, as policymakers and practitioners rush to take up value-added methods, researchers continue to debate their merits and how the existing models can be improved.

While value-added assessments are well past their infancy, noted Robert Lissitz, the director of the Maryland Assessment Research Center for Student Success at the University of Maryland, "the practical applications of value-added models are complex, difficult, and often controversial."

Fairer Measurement?

That hasn't stopped the momentum, which has gained steam in part because of the federal No Child Left Behind Act. The law requires states to test every student annually in reading



ACCEPTABLE

Access to the v
<http://ad.doubleclick.net/adi/ew/printspns>;
 listed in the category of: **W**
 has been denied because it may viol
 District's Acceptable Use Policy (regulations of thi

and mathematics in grades 3-8 and at least once in high school.

That mandate has opened up the possibility of tracking individual student growth from grade to grade in far more states, a prerequisite for value-added modeling. At the same time, concerns that the law's accountability provisions are unfair to schools has sent people scrambling for alternatives.

Sixteen state schools chiefs wrote to U.S. Secretary of Education Rod Paige earlier this year requesting the flexibility to use value-added or growth measures to meet the accountability requirements. States such as Ohio and Pennsylvania are now working to incorporate such models into their state accountability systems, joining existing ventures in Arizona, Florida, North Carolina, and Tennessee. And many other states, including Arkansas, California, Colorado, Louisiana and Minnesota, are considering adding value-added assessments.

One of the big attractions for educators is that value-added methods could provide a fairer way to measure school and teacher effectiveness than existing accountability systems.

The NCLB law, for example, judges schools primarily on the percentage of children who perform at the "proficient" level on state tests. Schools don't get credit for students who make lots of growth in a given year but still fail to reach the proficiency bar, or for advanced students who continue to progress.

Schools also are judged by comparing the performance of cohorts of students in successive years—for example, the performance of this year's 3rd graders vs. last year's 3rd graders—even though the two groups may be quite different. In contrast, value-added methods track the growth of each student in a school over time from the child's starting point.

Such methods also can provide schools with diagnostic information about the rate at which individual children are learning in particular subjects and classrooms, making it possible to target assistance to students or teachers or areas of the curriculum that need help.

In 2002, the Pennsylvania education department invited districts that were already testing in grades 3-8 to participate in a pilot value-added project, using the model that William L. Sanders developed for Tennessee in 1992. The plan is to take the project statewide next school year.

'A Great Diagnostic Tool'

The 4,500-student DuBois Area School District, about 100 miles from the Ohio border, signed up immediately.

"There are people who are really worried about this concept and want it to be perfect before we say yes," said Sharon Kirk, the superintendent of the district. She spoke last month at a conference in Columbus, Ohio, sponsored by Battelle for Kids, a nonprofit there that is working with about 80 Ohio districts on a value-added pilot using the Sanders method. "I can't imagine why we would not absolutely embrace information that is going to make us better."

One of the first things Ms. Kirk did was ask each principal to predict which group of students

his or her school was serving best. Daniel Hawkins, the principal of the DuBois Middle School, said he'd been confident the school was doing a fine job educating its most academically advanced students. When the data came back, it showed that in both math and reading, those students were making less progress over the course of a year than similarly high-performing students in other schools.

"I was wrong," he said, "obviously wrong."

Amy Short, an algebra teacher at the school, said educators realized they were spending too much time reviewing material at the start of each school year and needed to accelerate instruction.



At Ohio's DuBois Area Middle School, teachers such as Amy Short advance instruction based on "value added" findings.
—Paul A. Wilson for Education Week

The school set up four different levels of algebra and provided additional periods of math practice for students with the lowest math scores who also were falling behind their peers. Each week, teachers in the same grade and subject sat down to decide what they would teach in the coming week, and crafted nine-week assessments to track students' progress.

By 2003, DuBois Middle School students were demonstrating significantly more growth over the course of the year than similarly performing students elsewhere.

"I really like this because I think it's a great diagnostic tool for me," said Ms. Short, who uses the data on individual students to tell whether they need additional support or enrichment. "I thought I was teaching my kids better."

Research by Mr. Sanders and others in the field has found that the variability in effectiveness between classrooms within schools is three to four times greater than the variability across schools. Moreover, students assigned to highly effective teachers for several years running experience much more academic growth than students assigned to a string of particularly ineffective teachers, although the precise size of those effects and how long they persist are unclear.

Based on such findings, said Daniel Fallon, the chairman of the education division at the Carnegie Corporation of New York, people have come to recognize that the effects of good teaching "are profound and appear to be cumulative."

High Stakes?

Most people appear comfortable using value-added information as a powerful school improvement tool. The bigger question is whether states are ready to use such methods in high-stakes situations.

So far, the U.S. Department of Education has not permitted any state to use a value-added model to meet the requirements for adequate yearly progress under the No Child Left Behind law. And it's not certain the department has the authority to do so without changing the

statute.

Celia H. Sims, a special assistant in the department's office of elementary and secondary education, said at the time states submitted their accountability plans to the federal government, most didn't have in place the grades 3-8 testing or student-information systems that would permit them to track individual student gains over time.

"Value-added can certainly be used even right now as an additional academic indicator by the state," she noted, although no state has made that choice. In part, that's because additional academic indicators can only serve to increase the number of schools potentially identified for improvement under the federal law.

"States are still looking at how growth can fit within No Child Left Behind," Ms. Sims said. She does not know of any value-added model that specifies how much growth students must make each year, so that all students perform at the proficient level by 2013-14, as the law requires. "That's the non-negotiable," she said.

Researchers in at least three organizations—the Dover, N.H.-based Center for Assessment, the Portland, Ore.-based Northwest Evaluation Association and the Washington-based American Institutes for Research—have been working on models to combine value-added analyses with absolute measures of student performance, so that students would be on track to achieve proficiency by a specified point.

"This, to me, is a central issue with value-added," said Mitchell D. Chester, the assistant state superintendent for policy and accountability in the Ohio education department. By state law, the department must incorporate Mr. Sanders' value-added method into the accountability system by 2007-08. "How do you combine looking at progress with still trying to ensure youngsters in Ohio end up graduating with the skills and knowledge that they need to succeed beyond high schools?"

Some policymakers also are eager to use value-added models as part of teacher evaluation or compensation systems. But while many researchers and educators said value-added results might, eventually, be used as one component of such systems, they should not be the only measure.

"I think that really puts too much of a burden on value-added measures," said Henry I. Braun, a statistician with the Princeton, N.J.-based Educational Testing Service.

In general, such measures can distinguish between highly effective and ineffective teachers, based on the amount of growth their students make, researchers say, but they have a hard time distinguishing between the vast majority of teachers whose performance hovers around average.

Moreover, while value-added models can identify schools or teachers that appear more effective than others, they cannot identify what those teachers do that makes them more effective.

"In the earliest years of implementing a value-added assessment system, it's probably smart to lower the stakes," said Dale Ballou, a professor of education at Vanderbilt University in

Nashville, Tenn.

It's also unclear how such measures would work for teachers whose subjects are not measured by state tests.

Both the Ohio Federation of Teachers and the Ohio Education Association have supported the use of a growth measure as part of Ohio's accountability system.

"We felt there were a lot of hard-working people out there who were not getting adequate credit for moving kids along the way they do," said Debbie Tully, an official with the OFT, an affiliate of the American Federation of Teachers.

But while the union is "more than open" to using such measures as one component in teacher evaluation, Ms. Tully added, it's far too early to tell if it can be used as an evaluation tool.

'Under the Hood'

Yet for all the criticism of value-added methods, said Mr. Braun of the ETS, "we have to confront the logic behind the enthusiasm that we see out there in the world for value-added measures."

The key, he said, is for policymakers to "look under the hood," and not just take such measures at face value.

"I think the fact that people are taking this stuff seriously now is focusing people on the right questions," said Vanderbilt's Mr. Ballou.

While value-added models may eventually run up against insurmountable limitations, they're not there yet.

"All the other methods are also flawed," Mr. Ballou noted, "so if you're not going to use this one, what's the alternative?"

PHOTO: Data from a "value added" pilot made Principal Daniel Hawkins realize DuBois Area Middle School was depriving bright students.

—Paul A. Wilson for Education Week

Vol. 24, Issue 12, Pages 1,14-15