

Education Week

Published Online: May 20, 2008

Published in Print: May 21, 2008

Commentary

Fixing the Flaw in the ‘Growth Model’

And Helping Schools, States, and NCLB in the Process

By David P. Sokola, Howard M. Weinberg, Robert J. Andrzejewski, & Nancy A. Doorey

Almost from the day the federal No Child Left Behind Act became law in 2002, educators, parents, and others have been calling for more flexibility in the way it allows states to assess student achievement. Over the years, that call has only grown louder: Eighty-two percent of Americans in last year’s Phi Delta Kappa/Gallup poll said that they wanted schools rated on the improvement students make during the year, rather than on the percentage who meet the state standard at the end of the year.

Policymakers have heeded the message—or at least part of it. In 2005, the U.S. Department of Education allowed up to 10 states to pilot assessment systems based on “growth models” that measure students’ progress toward standards, in contrast to the pass-fail assessments originally required by NCLB. Last December, U.S. Secretary of Education Margaret Spellings opened the growth-model pilot to all eligible states, greatly expanding access not only to a more equitable accountability system, but also to a richer stream of information that schools and educators can use to address the needs of their students.

But, please, let’s hold our applause—we’re not there yet. For while the expanded pilot lets more states develop their own growth models, it still requires those models to incorporate the traditional grade-level assessments. As long as this language is in place, the growth-model pilot will be laboring under a fundamental flaw: using an instrument designed to give static information about a relatively narrow slice of the achievement spectrum to measure a much broader range of movement over time.

Why not craft the reauthorized NCLB to foster innovation and improvement in the field of assessment, rather than to prevent it?

This makes about as much sense as using a snapshot to describe a marathon: It gives some information about who’s at or very near a given point, but says nothing about those who are behind or ahead of that benchmark. In contrast, a computer-assisted, fully adaptive assessment, currently prohibited under NCLB, draws on questions from multiple grade levels, dynamically adjusting them based on each student’s responses, to capture a student’s current position anywhere along the continuum, as well as that student’s distance from the grade-level proficiency marker.

Fortunately, correcting the flaw is relatively easy: Simply change the legislation’s language to permit states to use fully adaptive assessments in their growth-model pilots, if they so choose. U.S. Reps. Tom Petri, R-Wis., and David Wu, D-Ore., introduced such legislation in the House

last fall. Now, we're urging Sen. Edward M. Kennedy, D-Mass., to make the change a priority as he works with President Bush to reauthorize the No Child Left Behind Act.

For those who like the federal law as it is, this change would have minimal impact. States still could use their current grade-level assessments to measure student growth, or they could even forgo the option of growth models entirely. But for those who believe that NCLB is a work in progress, permitting the use of fully adaptive assessments would enable the law to realize its goal of increasing and determining the pace of student achievement.

First, this change would give schools credit for growth that grade-level tests can't reliably detect. Most educators, for example, would give good marks to a school that moves a 5th grader from 2nd grade skills to a 4th grade level over the course of the school year. But grade-level tests, by definition, aren't constructed to record achievement outside a narrow band of expectations. So the farther above or below grade level the student starts out, the less information a grade-level test will provide, and the greater the margin of error in individual scores.

While this flaw is NCLB's, it is schools, educators, and students who are unfairly paying the price—particularly those in areas with the highest concentrations of low-income and minority students. How many of those schools facing the law's numerous penalties—including the ultimate penalty of restructuring—are, in fact, achieving the intent of the growth models: placing their students on track to reach higher academic standards within a few years? Sadly, under the law's current rules, we have no way of knowing. As a result, the growth models as currently approved are having little or no impact.

This limitation became clear when our state, Delaware, implemented its own growth-model pilot with 32,000 students in 47 schools. In the summer of 2007, the pilot partners commissioned the Center for Research in Educational Policy at the University of Memphis to compare official NCLB status and growth-model school ratings with simulated growth-model ratings that used both the regular state assessment to determine proficiency and the adaptive assessment to determine growth for students below the bar. The center's results clearly showed the fully adaptive tests more accurately measuring whether a student was on track to reach proficiency within four years.

Using fully adaptive assessments also would, at long last, enable states to turn NCLB's blunt-force, pass-fail results into much more nuanced, relevant, and timely information that teachers could use to improve their instruction. One common criticism of the law is that it pushes schools to teach to the middle—focusing on students within a point or two of proficiency to improve their percentages—while neglecting higher and lower performers. Having detailed information about each student generated by the adaptive tests would help them teach to each child's needs, and set meaningful individual growth targets within the framework of the standards.

Schools also could use this high-quality data to identify effective practices, evaluate and adjust their instruction, focus their resources, monitor trends within their systems, and provide parents with information they could use to support their children's learning. The Northwest Evaluation Association, which produced the proprietary adaptive assessment we used in our pilot, has published growth norms based on the past scores of more than 2.3 million students. So educators and parents could compare a student's rate of growth with typical rates for a given score and grade level. Northwest is one of several vendors of such assessments.

The test design that adaptive assessments make use of is not a new one—in fact, it has long proved its worth in entrance exams such as the SAT, the Graduate Record Examination, and the Law School Admission Test, as well as in numerous national certification exams. Critics of adaptive assessments rightly raise the issue of costs, including the cost of aligning the tests to

state standards. But our pilot also looked into some ways these costs can be mitigated. The most promising strategy builds on the recent trend among states of identifying “essential standards”—the core building blocks in each content area. Such an approach opens the doors to much broader collaboration and cost-sharing among consortia of states on their assessments and online reporting systems. Why not craft the reauthorized NCLB to foster innovation and improvement in the field of assessment, rather than to prevent it?

In establishing national accountability goals, the No Child Left Behind Act also set a high bar for itself: If we intend to rate or sanction schools, educators, and students based on their performance, the data we use to evaluate that performance also need to be accurate and useful. This simply won't happen as long as the law fails to acknowledge the fact that not all students start out at grade level—and that it can't address the significant minority who fall outside its own fairly rigid bounds.

Ironically, these students—both the high and the low performers within their grades—are the ones who stand to gain the most from the targeted instruction and other benefits made possible by a properly functioning assessment system. Making it right is an easy fix, and the payoff would be huge. Now it's up to our federal lawmakers to make that happen.

David P. Sokola, who has served in the Delaware Senate since 1990, is a board member of the Southern Regional Education Board and the immediate past vice chairman of the Education Commission of the States. Howard M. Weinberg is the executive director of the Delaware State Education Association. Robert J. Andrzejewski is the superintendent of the Red Clay Consolidated School District and represents the Delaware superintendents on the state assessment task force. Nancy A. Doorey coordinates the Delaware Growth Assessment Pilot and represents the Metropolitan Wilmington Urban League.

Vol. 27, Issue 38, Pages 26-27, 29