

Study Flags Drawbacks in Growth Models for AYP

By Sarah D. Spark

Amid battles over teacher quality and school restructuring, there's one thing everyone seems to want in the next version of the Elementary and Secondary Education Act: an accountability system that measures student growth.

Yet the results of the U.S. Department of Education's growth-model pilot program, whose **final evaluation** was released earlier this year, suggest lawmakers may have to do some heavy lifting to include growth in accountability. Not only do state growth models vary considerably, but they also play out in ways that can run counter to the aims of providing greater transparency and better accountability for all students, not just those "on the bubble," or just below passing rates for their state exams.

"It seems to me there is a serious disconnect between the rhetoric supporting growth models and the incentives and structures they end up creating," said Andrew D. Ho, an assistant professor at the Harvard Graduate School of Education and a co-author of the federal growth-model pilot program evaluation.

Daria Hall, the director of K-12 policy development for the Education Trust, a Washington research and advocacy group, and one of the original peer reviewers for state growth-model proposals, agreed that some of the rhetoric in favor the concept has not been supported by the data.

"While there are certainly students who are not proficient but are making big learning gains, there's not nearly enough of them and not nearly as many as folks hoped or assumed that there were," Ms. Hall said, "and that's a real problem."

Department's Plan

More than half of states are already using or developing their own growth models, and incorporating growth into the next federal accountability system has become one of the most often-requested changes to the ESEA, whose current edition, the No Child Left Behind Act, was signed into law in

[← Back to Story](#)



EDUCATE YOUR EDUCATORS.

SHARE EDUCATION WEEK WITH YOUR COLLEAGUES.

+ Click here for unlimited access.

2002. Proposals to use growth models have the support of 15 national education organizations, including groups representing state schools chiefs, legislatures, governors, and school boards, as well as the National Education Association and the American Federation of Teachers.

Growth-based accountability is also a centerpiece of the Education Department's vision for the ESEA reauthorization. Secretary of Education Arne Duncan **told** House education committee members at a hearing last month: "[W]e mean a system of accountability based on individual student growth—one that recognizes and rewards success and holds us all accountable for the quality of education we provide to every single student in America."

"This is a sea change from the current law—which simply allows every state to set an arbitrary bar for proficiency—and measures only whether students are above or below the bar," he added.

Growth models have gained popularity because supporters say they provide that more nuanced picture of how students are progressing academically and what schools contribute to their learning.

"Simply counting the percent proficient is not a very good way to evaluate a school," said Peter G. Goldschmidt, a senior researcher at the National Center for Research on Evaluation, Standards, and Student Testing at the University of California Los Angeles, who has studied growth models. "You want to see how schools are facilitating learning, for which you need to look at individual kids."

Former Education Secretary Margaret Spellings started allowing states to experiment with growth models in 2005, via a pilot initially limited to 10 states. Each state had to tie growth to the existing annual proficiency targets for math and reading under NCLB, rather than setting different expectations for students based on their backgrounds or their schools' characteristics.

Ohio an Outlier

The Education Department evaluated only the states in the original pilot: Alaska, Arizona, Arkansas, Delaware, Florida, Iowa, North Carolina, Ohio, and Tennessee. Of those, only Ohio has shown the growth model to make a big difference in the number of schools that made adequate yearly progress: More than twice as many Ohio schools made AYP via showing their students grew academically than by increasing the percentages of students who hit proficiency targets in 2007-08. Evaluators found, however, that Ohio uses a much more inclusive definition of students' being on track than other states. For the rest of the pilot states, growth alone accounted for a mere 4 percent of schools making AYP.

"It's not surprising then that the growth model didn't have much effect," Mr. Goldschmidt said. "There were a whole slew of adjustments ... and when you took all of these, the growth model really became just another adjustment to how you count the percentage [of students] proficient."

Aside from Delaware, states do not use growth as a primary accountability measure for all students, the evaluation shows. Instead, schools are judged first on the basic proficiency status of each student group, and then through the use of those other "adjustments," such as a confidence interval to correct for changes in group size from year to year, or the federal law's "safe harbor" provision, which credits a school for improving by 10 percent or more from the previous year.

The pilot evaluation found more than 83 percent of schools made AYP by one of those standard

measures, meaning growth didn't figure much into the accountability picture.

Moreover, growth models in most cases didn't hold schools accountable for high-achieving students found to be falling off track. Some states, like Colorado and Tennessee, require schools to base growth accountability only on the students who are on track to meet their ultimate target, regardless of their current performance. Most states base growth accountability only on the students who are not on grade level now, which can obscure future problems, Mr. Goldschmidt said.

'Rube Goldberg' Models

Damian W. Betebenner, an architect of the Colorado model and an associate at the National Center for the Improvement of Educational Assessment, in Dover, N.H., agreed: "In my original forays into this, the [status] accountability system was often indistinguishable from the growth model. It could be a Rube Goldberg machine that kind of led you to the 'yes AYP/no AYP' decision, but there was nothing else in it."

Colorado's growth model, which has become a model for at least 15 other states, reports growth and status measures for all students and is the only state to allow parents as well as educators to use its database to look at a student's predicted achievement over various time frames, to understand how quickly the student will need to advance.

"You have some students starting 500 miles away from the destination and expect them to get to the destination in two hours; we can calculate that out and know they aren't going to get to the destination by driving," Mr. Betebenner said. "So we need to look at other ways to get them to the destination, or we need to consider that it will take them more time to get there."

Yet the implementation differences between states may make it harder for policymakers to glean best practices to include growth measures in the next ESEA.

"Even if we all had the same growth model, we could very well end up in a case where one state sees vast differences [in the number of schools making progress] and other states do not," Mr. Ho said.

Growth models have become popular in spite of the pilot's lackluster results. The Education Department opened the pilot to all states in 2007, and by 2010, according to a [study](#) by the Council of Chief State School Officers, 17 states had implemented a growth model and another 13 were developing one.

Not Value-Added

While they can at times rely on the same testing and other data, state growth models differ from the "value added" models that have attracted attention as a tool for evaluating teachers.

State growth models have myriad permutations, but they fall into three basic categories:

- The trajectory model, used by Colorado and Tennessee, among other states, is what most people

RELATED BLOG



[Visit this blog.](#)

think of when they envision growth: It takes the gap between the student's base test score and the proficiency target, usually three or four years out, to calculate how much the student must progress each year. A student who isn't on grade level this year, but whose prior test scores show he or she will reach proficiency within the allowed time frame, would be considered on track.

- A few states, such as Delaware, use a transition matrix. Rather than using test-score gaps, it measures how students achieve on a matrix of performance benchmarks, such as moving from the below-basic to the basic level.
- States like Ohio use a regression model, a statistical formula that predicts a student's likely achievement by comparing his or her test scores over several years with those of a representative cohort of students and then projecting the result out to the proficiency target. It can feel counterintuitive to teachers, Mr. Ho said, because it gives no weight to increasing test scores. "It basically says, if you have high scores now, but most of your scores are low, [the high score] is an anomaly," he said.

Mr. Ho and other researchers have **found** the regression model, also called a projection model, is the most accurate at predicting which students will actually be proficient at the end of the time frame.

Yet that model can send a grim message to struggling students. "It's even harder to demonstrate progress than [with] a status model," Mr. Ho said.

"You have to score really, really high, sometimes almost impossibly high, to demonstrate progress, so this is actually raising the standards on the lowest students."

In the end, Mr. Ho said, "none of [the growth models] has the perfect combination of transparency and incentives and rhetoric."