New Tests Put States on Hot Seat as Scores Plunge

Added rigor welcomed, but lower scores sting

By Andrew Ujifusa

As states begin to demand more rigor on their high-stakes tests—and the tests evolve to incorporate revised academic standards—many officials are gambling that an initial wave of lower scores will give way to greater student achievement in the future.

Changes to statewide tests and subsequent plummeting scores sparked controversy and emergency action in Florida last month, and similar shock waves have been felt as Kentucky, Michigan, Texas, and Virginia remake their testing regimes.

The increasing expectations are in many cases a preview of challenges expected nationally when new, rigorous assessments based on the Common Core State Standards are administered by nearly all states starting in 2014-15. To date, 46 states have agreed to adopt the common-core standards in English/language arts and 45 in math, and two consortia with various member states are spending $360 million in federal money to develop common assessments for the new standards.

States have long endured criticism that their existing tests, aimed at moving the states toward 2014 proficiency levels in reading and mathematics demanded by the No Child Left Behind Act, lacked the rigor necessary to gauge how well students stack up against the demands of college and workforce readiness.

For many states implementing a new generation of tests, there will be a "shock" as test scores drop, a dynamic that typically occurs with each evolution of such assessments, said Kathy Christie, a vice president at the Denver-based Education Commission of the States.

"I would anticipate that the performance is going to be poor, if indeed you have a state where they're making a pretty big shift in expectations. ... The failure numbers are big right off" and start to recover, she said. "That's what I would expect here."

Florida Collision

The conflict between tougher standards on high-stakes exams and the challenge officials face in responding to lower test scores is still playing out in Florida.

In May, the state imposed a tougher scoring method on the Florida Comprehensive Assessment Test, or FCAT, writing exam for such items as grammar and punctuation. That led to sharp drops in passing rates.
Those writing-test decreases, including a plunge to 27 percent proficiency among 4th graders, from 81 percent last year—so alarmed state school officials and the public that the state school board lowered the cutoff score retroactively, which buffered the impact on schools' ratings. It did so even though board members already had expected a drop in passing rates.

Florida Education Commissioner Gerard Robinson admitted in comments after the results that officials, in changing the FCAT grading system, "did not give enough attention to communicating these basic expectations to our teachers" on what the tests would demand.

But state officials also noted that they are well aware that the new, common-core-based assessments are on the way and are trying to prepare for them. States that use common-core assessments won't be able to change the cutoff scores as Florida did with its FCAT.

**Kentucky Braced**

But in general, focusing on the scores from such newly revised tests misses the point, said Kentucky Education Commissioner Terry Holliday. States are reorienting their tests to measure college and career readiness instead of proficiency, he noted. In effect, they are "measuring something different," Mr. Holliday said.

The first state to adopt the common-core standards, Kentucky used a new set of assessments for the 2011-12 school year called **K-PREP** for grades 3-8. Students are also taking new end-of-course tests in high school. In addition, Kentucky now requires all 11th graders to take the ACT college-admissions exam.

Scores on the exams will drop, Mr. Holliday predicted, because students taking K-PREP now have to deal with longer, nonfiction reading passages, for example, and exhibit greater "technical fluency" in their comprehension skills. But Kentucky officials feel that the state tests match what the common core will demand since the state worked in developing K-

**Setting the Bar**

States are feeling the heat after adjusting the way student performance is judged on their high-stakes tests, causing student scores to plunge in some cases. Changes come as states gird for assessments based on the Common Core State Standards in 2014-15.

**Florida**

Scores on the state’s FCAT writing tests at various grade levels plunged dramatically this year after the scoring scale’s difficulty was increased by giving grammar and punctuation more weight. In response to the drop and public concern, the state board of education lowered the passing score.

**Kentucky**

The first state to officially adopt the common core, Kentucky has new assessments this year in both grades 3-8 and in high school. The high school tests are designed to count 20 percent toward a student’s final grades in their respective subject areas.

**Michigan**

In 2011, the state set new cutoff scores for both its Michigan Education Assessment Program (for grades 3-9) and Michigan Merit Exam for high schools based on revised standards that include the ACT college-entrance exam. The Merit Exam is designed to show that a student achieving that score or better should receive a B or higher in that subject’s entry-level college course.

**Texas**

After four years of work, new performance standards for 15 end-of-course exams on the State of Texas Assessments of Academic Readiness went into effect this year. The passing standards will increase over the next four years, and more than 1,000 school districts have delayed some requirements for a year.

**Virginia**

In both 2011 and 2012, the state
superintendent warned lawmakers that scores on the revised Standards of Learning tests likely would drop as a result of their increased difficulty; passing rates for high school students dropped by 27 percentage points in the fall 2011 administration of the new math tests.

SOURCES: Education Week; Associated Press

Many Southern states, in particular, have moved in the past decade to match their standards with those more closely aligned with the National Assessment of Educational Progress, known as "the nation's report card."

That move has helped states aiming to prepare for the common assessments, said Jeff Gagne, the director of education policies at the Southern Regional Education Board, which has 16 member states. Twelve member states, for example, now have high-stakes tests students must pass to graduate, he said.

"The idea of what's [to be] ... assessed won't be radically different from what's being assessed right now in our states," he said.

Michigan Gears Up

The Michigan Merit Exam for high school students (in addition to incorporating the ACT, like Kentucky) has been redesigned this year to show that students who score at or above the "proficient" level on a subject should be able to get at least a B on the freshman-level college exam in that subject at a public university in Michigan.

Cutoff scores for proficiency on the Michigan Educational Assessment Program, or MEAP, given in grades 3-9 each fall, also increased significantly this school year. Students needed to get 65 percent of answers correct to pass, instead of the previous standard of 39 percent.

Based on that new cutoff—not because of a change in the test itself—math proficiency rates statewide on MEAP in all grades dropped by roughly 35 percentage points from 2010 to 2011, when the new standards went into effect, said Joseph Martineau, the director of the office of educational assessment and accountability at the Michigan education department.

In terms of common-core readiness, "We feel like we are a little bit ahead of the game, and that will serve us well when we're going into this situation, when we're taking this test that is more rigorous," Mr. Martineau said.

To help the public understand the impact of the new cutoff scores, the department last November released information illustrating how much MEAP scores in each of the past four years would have dropped if the new standards and scoring had applied retroactively. For example, applying the new cutoff-scores to 2010 results, only 35 percent of 3rd graders would have scored proficient, instead of the 95 percent deemed so currently.
Trouble in Texas

Even in Texas, which is among just four states that have not adopted the common-core standards, rising rigor on statewide tests has led to worries from local school officials and to demands for change.

In the 2011-12 school year, high school freshmen took new end-of-course assessments in four subject areas, called the State of Texas Assessments of Academic Readiness, or STAAR. A student must pass a total of 15 tests in four subject areas (math, science, social studies, and English) in order to graduate, and the cutoff scores for proficiency will increase by about 14 percent in increments from 2012 to 2016.

But according to Austin-based Save Texas Schools, which champions education funding and limiting high-stakes standardized testing, more than 425 schools this year have adopted a resolution calling on the state legislature to re-examine STAAR's impact on classrooms.

"Essentially, the teachers are working on preparing kids without really knowing what's going to be assessed. ... I think there's been anxiety on the part of kids because of the increased rigor," said Ken Baliker, the president of the board of trustees of the 39,000-student Clear Creek Independent School District, the first district in Texas to adopt the resolution.

Fairness Issue Raised

Casey McCreary, the assistant executive director for education policy at the Texas Association of School Administrators, said it's unfair to simultaneously roll out new tests and load more pressure onto the tests by linking them to graduation and grades.

"It's going to be a disincentive for students that are struggling. ... Once parents understand how complicated it is for their child to graduate from high school, the backlash is only going to widen and get stronger," she said.

But there's also public sentiment for the view that new testing standards, like those on the FCAT this year, were overdue.

"Some of the things we've heard is people saying, 'So, spelling, grammar, and punctuation weren't [already] a part of the scores?' and more surprised at that," said Jaryn Emhof, a spokeswoman for the Tallahassee, Fla.-based Foundation for Excellence in Education.