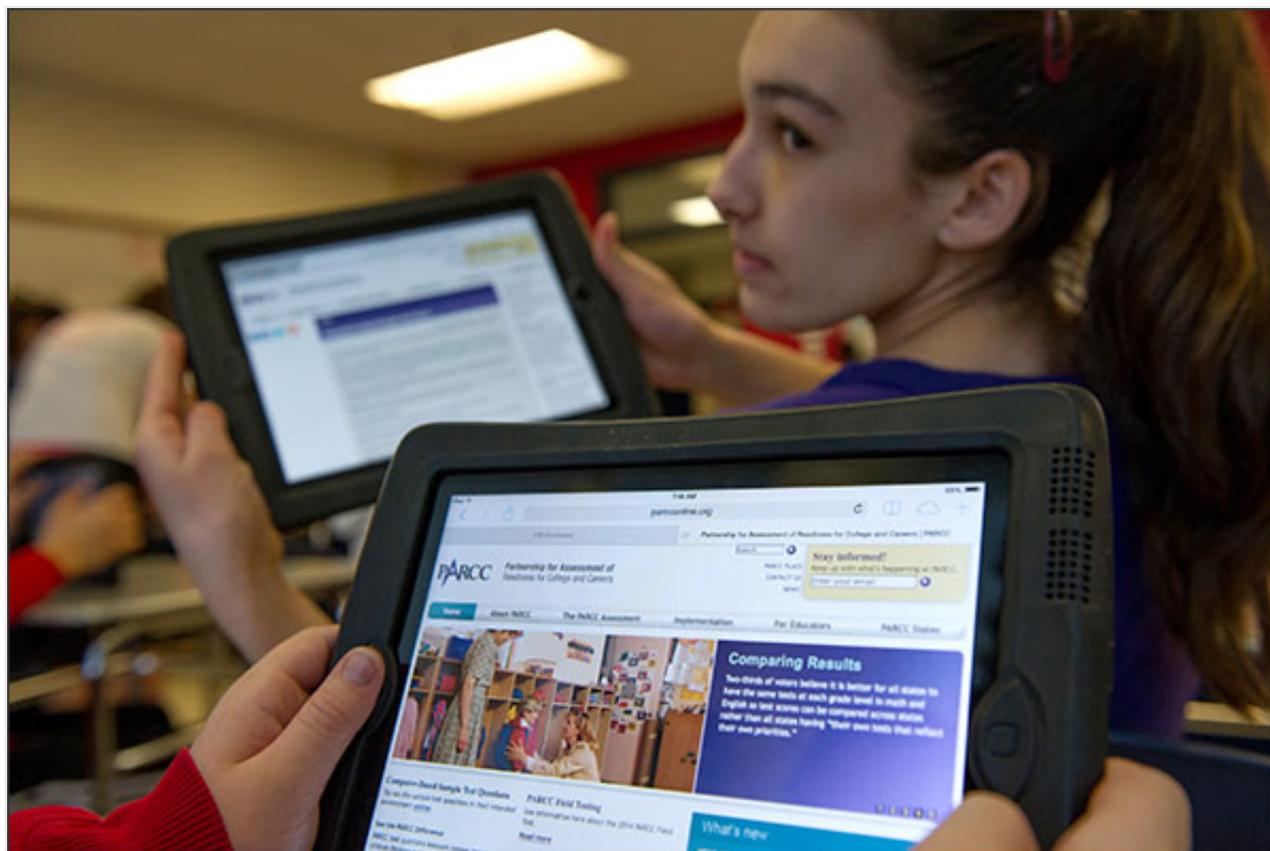


Field-Testing Set to Begin on Common-Core Exams

Published Online: March 21, 2014



Sydney Matai and other 7th graders at Marshall Simonds Middle School in Burlington, Mass., look at a PARCC practice test to give them some familiarity with the format before next week's field-testing of the computer-based assessments aligned with the common core.

—Gretchen Ertl for Education Week

By Catherine Gewertz

Next week marks a major milestone in an assessment project of unprecedented scope: the start of field-testing season for new, shared tests of a common set of academic standards.

Between March 24 and June 6, more than 4 million students in 36 states and the District of Columbia will take near-final versions of the tests in mathematics and English/language arts. Those exams—tied to the Common Core State Standards that all but a handful of states have adopted—were created by a bevy of vendors hired at the request of two groups of states: the **Smarter Balanced Assessment Consortium** and the **Partnership for Assessment of Readiness for College and Careers**, or PARCC.

“I don’t think a trial of this magnitude has been done anytime in the history of student testing in the U.S.,” said Keith Rust, a vice president at the Rockville, Md.-based Westat, where he oversees the sampling of schools and students for the National Assessment of Educational Progress, or NAEP.

The exercise won’t produce detailed, scaled scores of student performance; that part is still a year away. Instead, this spring’s field-testing is a crucial part of the assessments’ design stage, undertaken to see what works and what doesn’t. Questions like these are on test-makers’ minds: Will schools’ hardware and bandwidth be able to handle large-scale, computer-based testing? Do the tests work equally well on desktops, laptops, and tablets? Which items might confuse or overwhelm students?

Immense stakes are riding on the field tests. The federal government is watching closely to see how well its \$360 million investment—awarded in grants to the state consortia developing the exams—is paying off so far, especially because it has let more than a dozen states drop all or part of their current testing regimens in order to participate fully in the field tests.

States that pledged loyalty to the project need to see that they can rely on the tests, since those states are planning to base crucial decisions on them—such as how to evaluate schools, teachers, and students—within a year or two after the final tests are available in spring 2015.

School districts have made massive investments in technology to manage the

consortium tests, and have spent countless hours preparing their teachers, students, and parents for the new system—all on the faith that enduring the inevitable problems during the transition will pay off in a much better assessment than what they've been using. Amid a wave of anti-testing sentiment, many parents and activists are poised to seize on problems in field-testing as one more sign that large-scale testing is fundamentally misguided.

Combined, those elements create a combustible moment: An experiment deliberately designed to uncover weaknesses in a high-profile test takes place under intense public scrutiny.

“The consortia are going to have to be pretty confident they'll see minor glitches, but not major problems,” Mr. Rust said. “You wouldn't want to go into this on a wing and a prayer. If it goes badly wrong, it shakes people's confidence that it will be right the next time.”

In fact, just days before the planned March 18 start date for field-testing by Smarter Balanced, the organization took the major step of postponing the launch by one week to allow time for what Jacqueline King, a consortium spokeswoman, called some final “quality checking.” She said the delay was not about the test's content, but rather ensuring that all the important elements, including the software and accessibility features—such as read-aloud assistance for certain students with disabilities—were working together seamlessly.

A Representative Sample

There are some key differences between the field tests and the fully operational assessments that will be used in the spring of 2015. Length, for instance: Students will typically be involved in three to four hours of field-testing, less than half as long as what they'll face next spring.

In the real PARCC test, students will take both a multiple-choice, end-of-year

component and a more extended and complex performance-based section. On the field test, only 25 percent to 30 percent of students will take both pieces, and only in one subject, said Jeffrey Nellhaus, the director of assessment for PARCC. The rest will take either the end-of-year or performance segment.

The Smarter Balanced operational test in 2015 will be computer-adaptive—adjusting the difficulty of questions to the student’s skill level—but the field test, for the most part, will not be. A small number of students will get the adaptive version at the end of the field-testing window, said Jacqueline King, a consortium spokeswoman. That’s because test-makers will use the questions students answer earlier in the field test to calibrate the adaptivity of the test engine later in the field-testing window.

While some schools volunteered to participate in the field tests, most were chosen by their state or their state’s consortium as the multistate groups sought to build demographically representative samples of students. The result is a distribution of students taking the field tests that is wide nationally but not, in general, deep in individual schools.

The PARCC field tests, for instance, will involve about 10 percent of the students in the participating states and districts, but they are scattered across half the schools. That pattern is deliberate and beneficial, Mr. Nellhaus said.

“A more spread-out testing pattern,” he said, “means that you won’t get a clustering effect in the sampling” that could magnify the impact of anomalous conditions in any one place. “It also avoids a heavy impact on school life.”

Going All In

Most students are taking the field tests in only math or English/language arts; a subset will be tested in both subjects. Some states, however, such as California, Connecticut, Idaho, Montana, and South Dakota, have chosen to wade much deeper into the field-test exercise. They’re involving all—or nearly all—of their students. While that takes a greater

toll on schools' time and focus, leaders in those states decided that the payoff would justify the effort.

Those states were among the ones that obtained waivers from the U.S. Department of Education to cut back or eliminate their existing state tests to free up time to try the field tests. Because the new tests aren't final yet, the data they produce can't be used for accountability purposes, so the federal government **has agreed to let the waiver states hold their accountability ratings steady for another year.**

“We decided that it was a great opportunity for students to experience the test when it doesn't count, to see what's being asked of them,” said Deborah V.H. Sigman, the deputy state superintendent of education in California, where 95 percent of the students will answer Smarter Balanced field test items in both subjects, and the rest will take the test only in one content area.

“It's also a way for adults in the [school] building to think about what they need to do to optimize the experience for next year,” she said.

Some districts are doing more comprehensive field-testing than their states. The suburban system in Burlington, Mass., 15 miles west of Boston, chose to give the PARCC field test to every student in grades 3-11 in both subjects. Eric Conti, the superintendent of the 3,600-student district, said he thinks it's good for adults and students to experience something “as close to the real thing” as possible.

A federal waiver allows Massachusetts students participating in the PARCC field test to skip the state's regular testing under the Massachusetts Comprehensive Assessment System, although 10th graders will still have to take the MCAS to graduate from high



Dan Donoff, a Burlington, Mass., elementary school tech specialist, looks at a sample of the PARCC field test at the Marshall Simonds middle school in Burlington, Mass., last week, where teachers in the Burlington school district got a first look at the test and how it is formatted.

—Gretchen Ertl for Education Week

school.

Burlington was originally chosen by PARCC to do only the paper-and-pencil version of the field test, and only in some classrooms, in grades 3, 4, 8 and 10, Mr. Conti said. But he wanted to put his district's technological readiness to the test—it has a computer for every student—so he appealed to Massachusetts Commissioner of Education Mitchell D. Chester for permission to use the computer-based version of the field test with all children in tested grades, he said.

The district has made a deliberate research subject of itself, not only with PARCC, but with the Rennie Center for Education Research and Policy, a Cambridge, Mass.-based nonprofit. In partnership with the state teachers' union, the superintendents' association, and the state department of education, the Rennie Center will examine what happens in different field-testing scenarios in Burlington and in Revere, a small urban district north of Boston.

Burlington, for instance, will “livestream” the field test, so any loss of its network connections will interrupt the exam availability, Mr. Conti said. Revere, on the other hand, will “cache” the field test, downloading it and pumping it out locally. Burlington is trying the field test on varying devices, including iPads, Chromebooks, Mac desktops, and PCs, in a deliberate bid to see what works well and what doesn't.

“It makes no sense to show off technologically,” Mr. Conti said. “We could probably test all our kids in three days. Our network could handle it. Instead, it will be a three-week disruption.

“But the point is to see what happens,” he said. “As a superintendent, I plan 18 months in advance. When we do it live a year from now, it will impact my budget if we have to make changes. I'd rather know that sooner than later.”

Balancing Opposition, Potential

The Nashville, Tenn., school system illustrates both the promise and the risks districts

face when taking part in what the consortium test designers call “testing the test.” About 10 percent of the district’s 83,000 students will take the PARCC field test, either in math or in English/language arts.

Jessie Register, the district’s director of schools, said he thinks the experience will “take away the fear of the unknown” for teachers, students, and parents. It also complements the work the district has been doing to invest heavily in technological infrastructure and in training teachers to use technology to differentiate instruction, he said.

Since Nashville’s schools enroll one-third of Tennessee’s English-language learners, Mr. Register considers his district’s participation pivotal to ensuring that the PARCC test works well for students whose native language isn’t English.

“For our data to be included in how PARCC is going is important to influencing the design of the test,” he said.

But even as the Nashville schools inform a potentially better test, the district is treading on bumpy turf. Without a federal waiver for Tennessee, Nashville’s students will have to take both the PARCC field tests and the state’s regular assessments. And that likely won’t happen without some criticism from his community, Mr. Register said.

“We’re getting some pushback now about too much assessment,” he said. “We have to communicate very effectively with our parents and with our teachers to make sure this doesn’t become a negative.”

More than a few worries are shadowing the landscape as field-testing gets underway. Technological capacity is high on the list.



Members of the Burlington public school technology staff, including Tim Calvin, reflected at center, and Dennis Villano, reflected at bottom right, check the battery power on iPads that will be used for students taking the PARCC field tests in Burlington, Mass.

— Gretchen Ertl for Education Week

“We have 60 computers in one computer lab in our school. Our tech people are worried about our servers,” said Kristin Winder, a 6th grade teacher in Great Falls, Mont.

One district experienced such problems in the run-up to the PARCC field tests that it decided against participating. District sources said their faith was undermined by last-minute changes in test dates, student files uploaded but then lost, and other logistical and communications slip-ups.

“We simply couldn’t allow our system’s first experience with PARCC to be a negative one,” a district official said in a confidential email obtained by *Education Week*. “We believe it would have undermined our work and our staff. Students and parents deserve better.”

The complexity of mounting field tests on such a large scale is daunting. PARCC’s field-test-administration manual weighs in at 180 pages. The readiness exercise has spawned countless memos and staff meetings across states and districts as systems gear up for the field test. Educators have spent time trying practice tests with students, and administrators overseeing the coming exams have experimented with “training tests.”

“You can imagine the planning it takes to put something like this in place,” Ms. Sigman said of preparing for the Smarter Balanced field tests.

Some see big benefits in all that planning, as it provides a glimpse into how the common standards should inform instruction and a preview of the forthcoming tests. Others see those hours as a tragic mischanneling of education energy and resources.

“Schools are spending all this money trying to get wired and ready for PARCC and Smarter Balanced. And who’s getting that money? Corporations,” said Peggy Robertson, an Aurora, Colo., literacy coach who co-founded United Opt Out National, which seeks to eliminate high-stakes standardized tests. “The less money schools have, the more likely it is that they’ll fail. All of it is a setup for charter schools and the privatization of public education.”

Looking for Weak Spots

Teams from each consortium will be watching many aspects of field-testing closely to figure out what works well and what doesn't.

Questions of technology loom large: How many children can a given school test at one time? If a teacher is streaming video in her classroom while other children take the test down the hall, will it overload the system?

The teams are looking for many other outcomes as well. What kinds of answers does a given question elicit from a range of students? Test designers will have detailed student-level information—pegged to unique new identifiers to protect students' identities—to enable them to see if some questions stump subgroups of students, such as those in a given area of the country, English-language learners and students with disabilities, or those from certain racial or socioeconomic backgrounds. Do students who perform well on most parts of the field test consistently trip on some items?

Those kinds of observations will lead to a weeding-out or revision of questions, typically as many as 10 to 20 percent of the total, Ms. King of Smarter Balanced said.

Other research questions involve how to scale and score the tests. PARCC officials, for instance, will be considering whether to treat the end-of-year portion and the performance-task portion as separate exams, with separate scales and scores, or to combine them into “one big test,” Mr. Nellhaus said. And if they are combined, should the two pieces be weighted differently?

In the end, the two consortia are keenly aware that they're asking a lot of participating schools and districts: major time investments and schedule disruptions for what amounts to a research project to refine the test.

“This is why we do this,” Ms. King said. “To see what works and what doesn't.”

Even some of those most committed to the project are feeling trepidation. One district

official who described himself as “knee deep” in preparations said he is bracing for blowback from his staff and his parent community if even moderate problems arise with the test.

“I just hope it’s worth it in the end,” he said.

TESTING THE TESTS

The two state consortia that are designing tests for the Common Core State Standards begin field-testing their assessments this month.

FIELD-TESTING WINDOWS

PARCC

MARCH 24 - APRIL 11
for performance-task component; May 5 - June 6 for end-of-year component

MARCH 25 - JUNE 6

PROJECTED TESTING TIME FOR MOST STUDENTS

3 HOURS OR LESS (the operational test, in 2015, will be **8 TO 10 HOURS**)

3 TO 4½ HOURS (the operational test, in 2015, will be **7 TO 8½** hours)

NUMBER OF SCHOOLS PARTICIPATING

About **16,000**

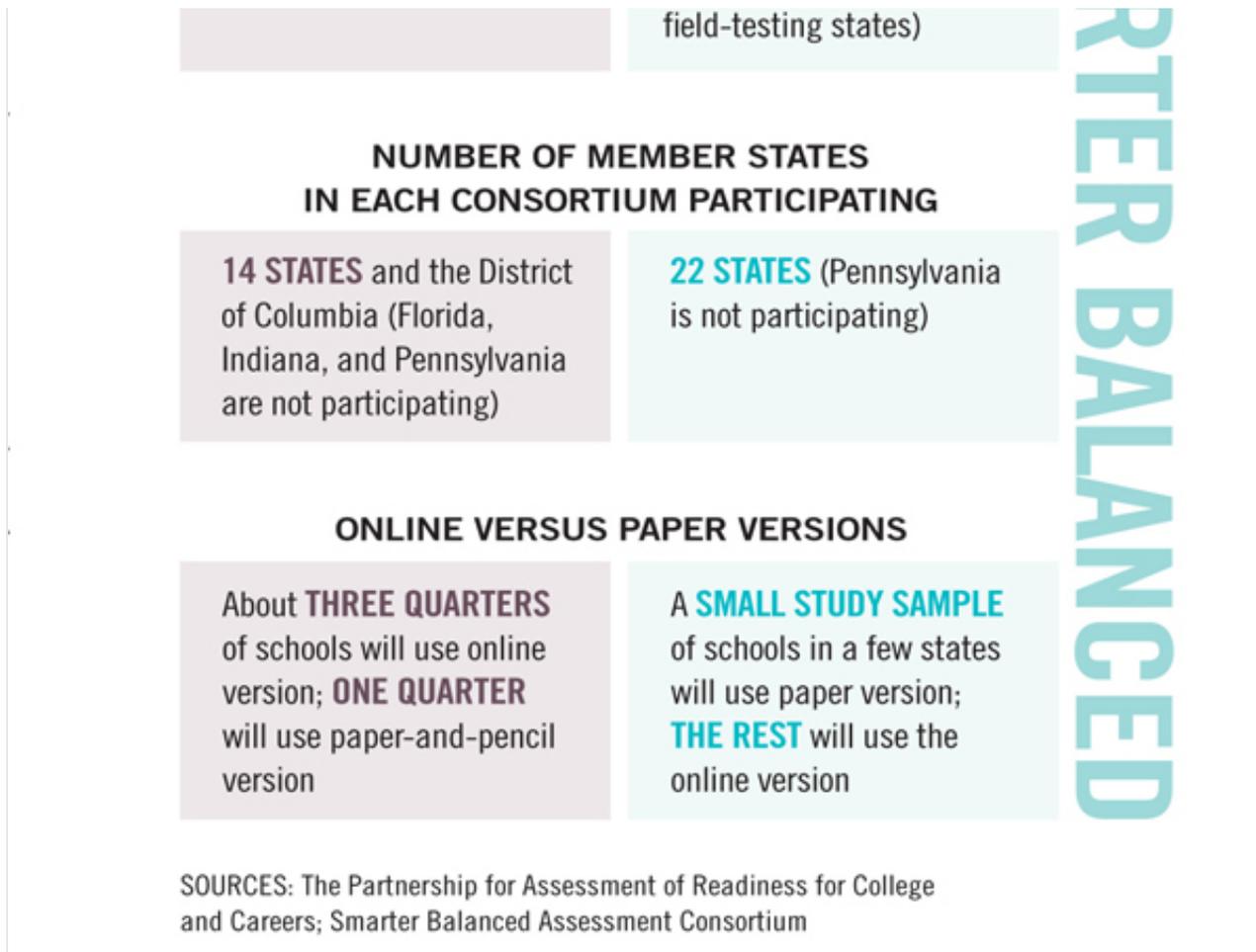
About **21,000**

NUMBER AND GRADE LEVEL OF STUDENTS PARTICIPATING

1 MILLION, grades 3-11 (of **10 MILLION** students in tested grades in PARCC field-testing states)

More than **3 MILLION**, grades 3-8 and 11 (of total **8.7 MILLION** students in tested grades in SBAC)

SMAI



Vol. 33, Issue 26

You must be logged in to leave a comment. [Login](#) | [Register](#)

[Back to Top](#) ▲