Editor’s Note: The common core has challenged test-makers and educators to craft assessments that measure the depth of knowledge expected by the standards. This Spotlight examines early results on common-core assessments, previews test items, and explores challenges still ahead for math and reading assessments.

TABLE OF CONTENTS:
1 Consortia Provide Preview of Common Tests
3 Transition to Online Testing Sparks Fears
5 Assessment Group Sets Accommodations Policy
7 D.C. Teachers Tally Results of Year’s Work
9 Challenges Seen in Testing Special Ed. Pupils on Common Core

COMMENTARY:
11 Why a Moratorium Won’t Work
12 Note to Congress: Fewer, Better Tests Can Boost Student Achievement

Consortia Provide Preview of Common Tests

By Catherine Gewertz

As teachers begin shaping lessons for the common standards, many are wondering how to prepare their students for tests that won’t be ready for at least two years. But sample items being drafted for those exams offer early ideas of what lies ahead.

Two large groups of states are using federal Race to the Top money to create a different approach. This item, from the Smarter Balanced Assessment Consortium, illustrates one of many examples of how two state testing consortia are trying to guide vendors as they design tests for the common standards.

Even if students don’t truly have a deep understanding of what two-fifths means, they are likely to choose Option B over the others because it looks like a more traditional way of representing fractions. Restructuring this problem into a multipart item offers a clearer sense of how deeply a student understands the concept of two-fifths.

A Different Approach

For numbers 1a-1d, state whether or not each figure has 2/5 of its whole shaded.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Option A</th>
<th>Option B</th>
<th>Option C</th>
<th>Option D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>1b</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>1c</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>1d</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

This item is more complex because students now have to look at each part separately and decide whether two-fifths can take different forms. The total number of ways to respond to this item is 16. Guessing the correct combination of responses is much less likely than for a traditional four-option selected-response item.

SOURCE: Adapted from Smarter Balanced Assessment Consortium
new suites of exams for the Common Core State Standards. Those consortia have recently begun work with private vendors to develop items—questions and tasks—for the tests. But each group has produced a range of sample test items to help those vendors get an idea of what the states want, and experts say they offer valuable insight into the tests that are expected to emerge in 2014-15.

“What we are starting to see here are tests that really get at a deeper understanding on the part of students, not just superficial knowledge," said Robert L. Linn, an assessment expert and professor emeritus of education at the University of Colorado at Boulder who reviewed a sampling of the consortia’s materials. “But unless students are really prepared for them, it’s going to be a huge challenge.”

Mr. Linn predicted that even with sample items to guide them, vendors will find it tough to develop tasks and questions that fully reflect the aims of the two state groups, the Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for Colleges and Careers, or PARCC.

“Where the real difficulty comes up is when you actually develop the items,” he said. “It will be a challenge for vendors to come up with items that meet these specifications. They are used to writing items for state tests that do not get at this depth of knowledge.”

Sample items for the English/language arts exams, crafted by work groups of the Smarter Balanced consortium in conjunction with test-makers ETS and Measured Progress, offer a glimpse of what that group has in mind. A few were presented at a meeting of the American Federation of Teachers in Baltimore this summer, and others have been circulating among states for feedback.

One selected-response item asks 5th graders to read an article about how scientists track bird migration and to identify the two paragraphs that contain the author’s opinions on the topic. The question taps key skills required in the common standards, such as comprehending “content rich” non-fiction and citing textual evidence for an argument.

A constructed-response item for 11th graders asks them to read excerpts from an 1872 speech by women’s rights activist Susan B. Anthony and the “Second Treatise of Civil Government” by English philosopher John Locke, published in 1690. They must identify the ideas common to both pieces and discuss how Locke's ideas support Anthony's arguments, citing evidence from each to support their interpretations.

One of the aspects of the consortia’s work that represents perhaps the greatest departure from current state testing practice is the inclusion of performance tasks, which engage students in more complex, prolonged exercises. Smarter Balanced’s policy director, Susan Gendron, a former commissioner of education in Maine, offered an example of that approach, too, at the Baltimore meeting.

“The sample task, scheduled to take 105 minutes, asks 6th graders to read an interview with a teenager who started a charity to help Peruvian orphans. It directs them to articles and videos on specified Web pages to learn more about other young people who devote themselves to helping those in need.

The students answer constructed-response questions that require them to describe what they’ve learned, analyze the meanings of key words, and discuss how they evaluated the reliability of their Web resources. They must research and present a five-minute speech about a “young wonder” of their choice, complete with audio-sual representations.

The fact that all students will now be asked to dive into complex performance tasks is a marked departure from current practice, according to Kris Kaase, who oversaw state assessments in Mississippi from 2002 to 2010.

“These are very high-level kinds of tasks students are going to be asked to do,” he said. “It is significant that all kids are being asked to do this sort of thing. Some students might have been tested in this way, but not all students.”

No Guesswork

Sample math items being circulated by the consortia attempt to plumb not only students’ procedural skills, but their conceptual understanding. Item specifications produced by Smarter Balanced, for instance, include, among other things, an alternative to a typical way of gauging elementary students’ grasp of fractions.

The more traditional approach offers students four graphical depictions of the fraction two-thirds and asks which is correct. The other approach offers the same four graphic models, but asks students to say whether each one has two-fifths of its whole shaded. The latter, according to the consortium’s notes, makes it harder to guess a right answer and requires an understanding of the different forms that fraction can take.

“To perform well on these kinds of assessment items, just having good test-taking skills will not be enough of an edge to perform well,” said Mr. Kaase, who now runs a Jackson-based consulting company that works with states and districts on testing, curriculum, and accountability issues.

“One of the questions is, can we make stronger claims [about students’ skills] from a differently designed item,” he said. “Because of the way these items are being constructed, it seems likely we would be able to make those stronger claims.”

Building Fluency

A sample math performance task by Smarter Balanced asks 6th graders to figure out what they need to build a community garden to a given set of specifications for $450.

During two test sessions totaling up to two hours, students would have to calculate many figures, including the perimeter, surface areas, and volume of each section of the garden, and make a sketch based on their calculations. They must figure out how much wood and soil are needed and how many tomato and carrot plants to buy, given their cost, the garden’s size, and each plant’s need for space. Finally, they must show how their project will stay within its allotted budget.

Materials developed by PARCC, too, illustrate for vendors item types that require a grasp of the topic, said Mr. Kaase. One, for instance, asks 4th graders to plot the following numbers along a number line: 2, 5/4, 3x1/2, 3/4+3/4, and 2-1/10.

“You have to understand the meaning of the numbers and how they relate in order to answer this,” Mr. Kaase said.

Chuck Pack, the chairman of the math department at Tahlequah High School in Oklahoma, which belongs to the PARCC consortium, said he is pleased with the sample items he sees being developed by the consortium.

“They made me hopeful,” said Mr. Pack, who is a teacher-leader for PARCC, helping colleagues deepen their knowledge of the group’s work. “I was a little concerned at first blush, because they’re really complex. But they’re good math problems. They’re above the level of what we’re currently doing, but they’re attainable.”

He pointed to one illustrative example in PARCC’s materials that tries to gauge students’ fluency in division and multiplication. It offers five equations, such as 54÷9=24÷6, and asks 3rd graders to specify whether each is true or false.

“I like that it does multiple assessments in one item,” he said. “It asks kids to work each of those problems easily and be comfortable with it, which is what fluency is.”

PARCC expects to release sample items in English/language arts and math later this month, including prototypes developed
under contract with the Institute for Learning at the University of Pittsburgh and the Charles A. Dana Center at the University of Texas at Austin.

**Time Crunch**

The state consortia have big challenges on their hands, however, and time is among the biggest, experts said. Mr. Kaase recalled that when he was Mississippi’s assessment director, it took at least three years to produce a new version of a state test.

In the case of the common assessments, he said, “we’re not talking about a new version of an existing assessment. We’re talking about a revolutionary kind of change,” not only with different kinds of items, but also computer-based or computer-adaptive technology. “What they are trying to do with these items pushes the bounds,” he said. “To get this done in the amount of time they have is going to be a challenge.”

Many milestones lie ahead before the consortia can deliver fully populated banks of test items. In the coming months, both groups will conduct sessions in which items are tried out with students and their feedback is obtained. Smarter Balanced will train teachers as item writers and has prepared a bank of training materials to assist in that work. Both consortia will conduct trials next year before full-fledged field tests in spring 2014.

Even as sample items are crafted to help guide vendors on item-writing, the consortium and their partners caution that the item-development process is lengthy and full of revisions.

In a presentation to the National Assessment Governing Board in Washington this month, Jeffrey Nellhaus, PARCC’s assessment director, said he was acutely aware that the “field is hungry to see” how the goals of the common standards will be “made manifest” in assessment items, and is eager to examine the prototype items the consortium will get from the two research universities.

“It’s a real challenge to assess the common core in assessment items,” he said. “We’ve been working with national experts. ... But it may take some iterations in our development process.”

As officials from the Dana Center cautioned in an overview of the PARCC project, “Prototyping is for learning, and it can be messy.”

---

**Transition to Online Testing Sparks Fears**

By Catherine Gewertz

When tens of millions of schoolchildren sit down at computers to take new common assessments in spring 2015, many of their peers will be taking similar tests the old-fashioned way, with paper and pencil, raising questions about the comparability of results—as well as educational equity—on an unprecedented scale.

Both state consortia that are designing tests for the Common Core State Standards are building computer-based assessments, but they will offer paper-and-pencil versions as well, as states transition fully to online testing. The Smarter Balanced Assessment Consortium plans to run the two simultaneous “modes” of testing for three years. The Partnership for Assessment of Readiness for College and Careers, or PARCC, will do so for at least one year.

In order to rely on the results, however, the consortia must show that the paper and computer modes of the tests in English/language arts and mathematics measure the same things.

The prospect of establishing such comparability between two versions of a test isn’t new. States have long used established statistical and psychometric practices to do so when they update their paper-and-pencil tests, for instance, or when they transition from paper-based tests to computer assessments. But the challenge before the two consortia ups the ante by hanging the validity of far more children’s test scores on the “linking” or “equating” process conducted by each group.

“In the assessment profession, we need to be able to back up claims we make about students’ and schools’ performance. Any threat to validity is a threat to those interpretations,” said Richard Patz, the chief measurement officer at ACT Inc., which is conducting comparability studies of its own as the Iowa City, Iowa-based company introduces a digital version of its college-entrance exam.

Thorny questions have arisen, too, about whether children who take the paper-and-pencil version of the consortia tests will be at a disadvantage—or perhaps have an edge—compared with their peers who take the computer version.

Could children in high-poverty areas, where technological readiness will likely be lower, lose something valuable by not interacting with the new tests’ technologically enhanced items, such as drawing and drag-and-drop functions? Would they actually benefit by sticking with paper exams if they are more comfortable taking tests in that mode?

**Mixed Landscape**

Consortia leaders say they are confident that comparability and equity questions will be fully addressed by the time the tests make their debut in 2015.

“It’s something we need to do carefully, and we intend to do it carefully,” said the executive director of the 25-state Smarter Balanced group, Joe Willhoft, who oversaw such studies as the assessment director in Washington state.

Jeffrey Nellhaus, the testing director for PARCC, which includes 18 states and the District of Columbia, said the group’s test designers are “very sensitive” to comparability questions and are planning studies to answer them.

About 40 million students attend school in the states that belong to the two consortia. But much is still unknown about how many will take paper tests in 2015 and how many will use a computer. Even rough feedback, however, shows a strong likelihood that large swaths of students will be picking up their No. 2 pencils.

Survey data collected in July by Smarter Balanced—also more of an approximation than a full accounting—show a wide range of technological readiness.

Oregon, long a leader in online assessment, reported that all its districts were capable of giving tests online, while only 45
percent of California’s districts did likewise. For PARCC, Mr. Nellhaus ventured a guess of a 50-50 split, but emphasized that data on districts’ and schools’ readiness are far from complete.

The consortia will not decide who takes the paper-and-pencil version of the test and who takes the computer version, officials said. That will be up to states, and in some cases, individual districts or schools.

Ideally, test results are “indifferent” to the mode in which the test is given, said Henry Braun, a longtime researcher with Princeton, N.J.-based test-maker Educational Testing Service and now an expert in educational evaluation and measurement at Boston College. If the mode of administration helps or hampers some students, the results are distorted, he said.

**Differences in Format**

Assessment experts say it’s much easier to establish comparability when two tests are similar in format, such as a multiple-choice test on paper that becomes a multiple-choice test on the computer. But even then, comparability issues can arise.

A student who must read a test passage in order to answer a multiple-choice question, for example, might be able to read the entire passage on one page of the paper test, but on the computer, she must scroll up and down to do so. Such shifts can affect the performance of some students, said a longtime assessment expert at a major testing company. (Like most experts interviewed for this story, he agreed to speak only if his name was withheld because of his employer’s contracts with the assessment consortia.)

Comparability challenges deepen when tests differ significantly in format, experts said. In the case of the two state consortia, their computer-based exams—with technology-enhanced items such as interactivity and animation, and longer, more complex performance tasks—will be able to represent ideas in ways that the paper versions cannot, so establishing comparability between the two will be tougher.

“When an assessment has types of items only available in one mode, it creates a greater challenge for establishing comparability, but it’s a familiar one and it’s generally manageable one,” said ACT’s Mr. Patz.

The other expert, however, said that while the consortia’s comparability challenge is “not a fatal problem, it needs to be thoughtfully negotiated and represented to anyone who will use those test scores.”

That source said it’s not possible to measure everything in the paper-and-pencil version that can be measured in the computer-based version.

“In the technical sense of ‘comparable,’ the two might not be comparable,” he said. “If you were successful in measuring the same things, which would be a stretch if the computer-based version’s items are truly innovative, it could well be the case that one [test] could be harder or easier than the other because of how the items are presented.”

**Writing From Scratch**

Assessment specialists outlined various ways to establish comparability between the paper and computer versions of a test. One is to use a set of common items in both, so test designers can compare student performance on those items in the two modes. Another is to randomly assign students to take one or the other mode of the test. Better yet, a study group of students can be selected to take both the paper and computer versions. Consortium officials said such methods are being planned or considered for field tests next spring.

Testing experts also said it’s best to create assessment questions from scratch for the paper-based assessment, rather than building paper versions of test items originally designed for the computer.

“You can’t replicate the interactivity of the computer environment on paper,” said one testing expert. “You need to build alternate forms of the test that measure the same standards [on paper].”

Mr. Willhoft from Smarter Balanced said that his group is adapting items written for the online environment to paper. Mr. Nellhaus from PARCC said its developers are writing paper items from scratch to use in place of technology-enhanced items on the computer, but more traditional item types can be used in both modes.

PARCC’s field test next spring will include paper-based as well as computer-based exams, Mr. Nellhaus said. The Smarter Balanced field test will include paper forms only for a small group of students, to study comparability, Mr. Willhoft said. “There’s no denying that there will be some items that will be difficult to translate into the paper environment,” said Mr. Willhoft. One of the consortium’s math items, for instance, asks students to click on images of a cylindrical shape and a rectangular one in an exercise about volume. “But there’s nothing inherent in a given standard that requires a certain kind of interactive item,” he said. “You can measure the same standard in different ways.”

Smarter Balanced faces an extra layer of complexity in comparability because its test is computer-adaptive, meaning it adjusts questions to the test-taker’s skill level.

“With an adaptive test, you see right away what questions a kid needs,” said Lauress L. Wise, a principal scientist with the Monterey, Calif.-based Human Resources Research Organization, which has performed quality assurance and evaluation on testing systems such as the National Assessment of Educational Progress. “With paper and pencil, you’d have to offer a lot more questions—a longer test—to make it comparable to that. If you can’t do that, you won’t be measuring the end points [of achievement] as well.”

Mr. Willhoft acknowledged that the paper version of the Smarter Balanced test will be “less precise, with a larger measurement error” at those points in the spectrum.

In seeking comparability, a key consideration is what kinds of conclusions will be drawn from the scores on the two types of tests, said Mr. Wise. The degree of comparability takes on added significance when high-stakes decisions are based on the results, he said.

“If this were a graduation test, and some kids were getting denied diplomas because they took one form or another, you could make a plausible argument why there could be a lawsuit,” Mr. Wise said. “That could get sticky.”

**Quality of Tasks**

The fact that paper-and-pencil tests might be more widely used in lower-income areas is something that officials at the Education Trust, which advocates school improvement for disadvantaged students, are keeping an eye on. But those potential questions of equity revolve more around the quality of the assessment—and the teaching that goes with it—than about the mode of the test, they say.

Christina Theokas, the organization’s director of research, said she worries that if the paper test is less complex and instructionally rich than the computer version, classroom instruction could mirror that.

But students aren’t necessarily at a disadvantage just by taking a paper-and-pencil test, said Sonja Brookins Santelises, the Education Trust’s vice president of K-12 policy and practice. Top-notch paper tests such as NAEP and Massachusetts’ statewide exams demonstrate that, she said. The important thing to watch is not the mode in which a test is administered, Ms. Santelises said, but the quality of the task and how well students are prepared for it.

“You can do a rudimentary task on a computer and have it not be beneficial, and you can have a paper-and-pencil task that’s instructionally rigorous and very beneficial,” she said. “Are students going to have access to the kind of experiences and curriculum that prepare them for those kinds of tasks? Are teachers being prepared and supported to do that?”

Ms. Santelises added: “We need to stay focused on the teaching and learning, rather than on whether we have the right technology to give a test.”
Assessment Group Sets Accommodations Policy

By Lesli A. Maxwell and Christina A. Samuels

Native-language translations of test directions and math items, as well as a read-aloud option for middle and high school students, are among the accommodations that will be allowed on a limited basis under a policy adopted this week by a group crafting common-core-aligned tests in reading and math for 24 states.

The Smarter Balanced Assessment Consortium’s governing board voted unanimously on “usability, accessibility, and accommodations” guidelines that outline the kinds of testing supports and tools that will be made available to all students, and particularly those with disabilities and English-language learners. The issue of reading text passages aloud on English/language arts tests drew the most debate, with some disability advocacy groups arguing that test-makers should limit read-aloud supports only in situations where students are being tested on their decoding skills. Others argue that listening comprehension is not the same as reading and, thus, not a valid gauge of students’ reading skills.

The policy attempts to create a middle ground. Students in grades 3-5, a stage when learning to decode text is central to the curriculum, will not have the option to have text passages read to them on the English/language arts portion of the test. But students in middle and high school will be able to use a read-aloud accommodation and have the test be considered valid if their need for such a support is documented in an education plan.

The policy also makes a distinction between text passages and test questions and directions. All students will have the option to have test questions and directions read aloud to them if a school team decides they need that support. Magda Chia, Smarter Balanced’s director of support for underrepresented students, said the board’s decision was based on expert advice that reading text passages aloud would, for younger students, fundamentally alter the skills the test is measuring.

An earlier draft of the guidelines allowed elementary students who are blind and still learning Braille to have text passages read aloud, but the consortium was told that it could not legally single out one group of students for support while denying the same accommodations to others, Ms. Chia said.

States do have the option to offer the read-aloud accommodation even if doing so goes against the guidelines, but in that case, the test won’t be counted as valid, she said.

Categories of Help

The policy manual includes three broad categories of testing supports for students:

• “Universal tools” are those that will be available to all students. Many are built into the technology platform students will use to take their tests and include English dictionaries, highlighters, spell-check programs, and breaks from testing.

• “Designated supports” are features that will be available to any student who can benefit from them, based on a teacher or school team’s discretion. Those supports would have to be identified for students ahead of any test administration. All the language supports available for English-learners fall into that category.

• “Documented accommodations” are for students whose individualized education or Section 504 plan calls for their use. That would include reading aloud text passages for middle and high school students, Braille, or closed captioning for hearing-impaired or deaf students.

Not every English-learner in the Smarter Balanced states will have access to the full range of available language supports. Member states with laws and regulations that restrict or prohibit the use of languages other than English to teach or assess ELLs do not have to offer such translation options for test-takers.

California, a Smarter Balanced state that is home to about 1.6 million English-learners, will offer the full array of language supports, said Deborah Sigman, the state’s deputy superintendent of education and a co-chair of Smarter Balanced’s executive committee. The state has offered a similar lineup of language supports to English-learners on its content assessments.

“At this point, we don’t see any conflicts with current law, but obviously, we will review as we begin our implementation of the Smarter Balanced assessments,” Ms. Sigman wrote in an email.

Translation Options

Among the language supports that ELLs may encounter on the Smarter Balanced math tests: translated test directions; translated glossaries for select words in math test items that help students understand specific terms; and complete or “stacked” translations of test items that will appear above the original English item. All of those translation tools will be embedded in a test’s digital platform and can be turned on or off by test administrators. In addition, ELLs may also use nondigital translated glossaries to understand selected
math terms on some test items. For the English/language arts assessment, Englishlearners may have access to a bilingual or dual-language word-to-word dictionary for the extended writing portions of an ELA performance task.

All the translation supports will be available in Spanish, the predominant native language of students who are English-language learners in Smarter Balanced states. But the test consortium will also offer translations in other languages, as determined by member states’ needs. So far, it has committed to providing the glossary translation tools in Vietnamese and Arabic, the next two most widely used languages in those states.

Gabriela Uro, the director of English-language-learner policy and research at the Council of the Great City Schools, said the language supports will help even the playing field for English-learners to demonstrate what they have learned.

And if scores for English-learners in states that offer the full array of language supports are better than for students who don’t get the tools, “it might open the door” for states with more restrictions around the language of instruction to think about how they can also offer the tools without violating their laws, Ms. Uro explained.

The other big group of states designing assessments—the Partnership for the Assessment of Readiness for College and Careers, or PARCC—approved its first edition of testing supports for ELLs and students with disabilities in June.

PARCC, which received hundreds of comments on its proposed accommodations policy—many relating specifically to read-aloud accommodations—decided in June that students with persistent text-decoding difficulties can have text passages read to them, but a notation will be made on their score report indicating that no claims can be made about the student’s ability to demonstrate foundational print skills.

Policy ‘Disappointing’

Timothy Shanahan, a professor of urban education at the University of Illinois at Chicago and a former president of the International Reading Association, offered a scathing assessment of PARCC’s read-aloud accommodations policy in his blog, saying it was an example of “dopey doings in the annals of testing.”

The Smarter Balanced approach to read-aloud accommodations is a better option, he said, though he would rather not see its use for text passages.

“The purpose of the test is to find out how well these kids can read,” Mr. Shanahan said. “If for whatever reason they can’t do it, that’s what we want to know.”

Lindsay E. Jones, the director of public policy and advocacy for the National Center for Learning Disabilities in New York, said she was “stunned” at Smarter Balanced’s blanket prohibition of read-aloud accommodations for students at the elementary level.

“Smarter Balanced has been inclusive in the discussions up to this point. That’s why I think this is so disappointing,” Ms. Jones said. She acknowledged concerns that read-aloud accommodations could mask inadequate instruction, but said the data that states will be able to collect through those tests will help pinpoint areas of concern or potential overuse.

“This conversation is not over,” she said. “We view this as the first step in what will hopefully be a long dialogue.”

The supports will be field-tested with student-test-takers next spring. The full rollout of the Smarter Balanced exams will be in the 2014-15 school year.
D.C. Teachers Tally Results of Year’s Work

By Catherine Gewertz
Washington

Staring at multicolored rows of names and numbers on a laptop screen, Dowan McNair-Lee is searching for clues to how well she taught her students.

The 2012-13 school year was a difficult journey, as the English/language arts teacher tried to move her challenging and varied group of 8th graders to mastery of the Common Core State Standards. Now, two weeks before the 2013-14 year begins, she scrolls through year-end test scores that deliver part of the verdict on her success.

The data analysis caps the second year of an unusually aggressive and comprehensive campaign to put the English/language arts curriculum into place at the Stu- hart-Hobson Middle School. Trying to ignite promising results, it can’t reach every student subgroup, right down to the individual academic standards themselves. What they find will help shape this year’s teaching and instructional coaching at the school level, and curriculum resources and professional development at the district level.

Scanning the rows of data, color-coded by achievement level, brings a roller coaster of reactions. Ms. McNair-Lee claps and beams when she notices a student who moved from the “basic” level of performance to “proficient.” “High fives!” she exclaims, raising one palm in the air. She applauds and smiles again when she scans a few rows down and sees another success story: a girl who had been high on the teacher’s radar because of her behavior and academic problems moved from proficient to “advanced.”

Only a moment later, Ms. McNair-Lee frowns and shakes her head. On her computer screen, she sees that two students who scored basic in 2012 slipped to “below basic” in 2013. One of them is Mikel Robinson, who seesawed academically all year long. In the end, he eluded her reach, leaving her wistful for a hug.

“Tears, I hate it, but there’s nothing I can do about it now. It’s over,” she says, softly. “Some of them, I sent them out well. And some of them, like Mikel, I keep wondering what more I could have done.”

Those emotional ups and downs permeate the mid-August dive into the test-score data by Ms. McNair-Lee and her colleagues from Stuart-Hobson Middle School. Trying to ignore the stuffy classroom heat and the jackhammers pounding outside, the educators spend hours bent over laptops and printouts, parsing which parts of their instruction went well last school year and which didn’t.

In classrooms all over the District of Columbia, the work of this small team is replicated as staff members from 111 schools prepare for the new year. They’re analyzing student performance on the DC CAS, the school system’s end-of-year test, by grade level, subject, student subgroup, right down to the individual academic standards themselves.

This isn’t episodic success, it’s systemic success. We saw growth at every single grade, in every ward, all the subgroups. That tells me that it’s not just dependent on the quality of the principal in one particular school. This work is landing, and landing consistently across the board.”

KAYA HENDERSON
Schools Chancellor, District of Columbia

As the District of Columbia has seen, even as a vast new push to make change can produce promising results, it can’t reach every student and teacher with the support they need.

In classrooms all over the District of Columbia, the work of this small team is replicated as staff members from 111 schools prepare for the new year. They’re analyzing student performance on the DC CAS, the school system’s end-of-year test, by grade level, subject, student subgroup, right down to the individual academic standards themselves.

Principal Dawn Clemens took the microphone and said: “By the power vested in me, I pronounce each and every one of you a high school student.” The room erupted in cheers and camera flashes, then emptied, slowly, into hallway gridlock. Without a tissue, Ms. McNair-Lee kept wiping her cheeks as swarms of girls locked her in group hugs.
On the front steps of the school, Mikel found a moment for a farewell hug from his English teacher. Then he turned his back to the school to join his family, and they drifted off down the sidewalk in the fading evening light.

**Making Headway**

Six weeks later, at a middle school a few miles away, districtwide test scores were announced. At a celebratory event featuring District of Columbia Mayor Vincent C. Gray and Schools Chancellor Kaya Henderson, the district school lauded its progress.

While city schools are far from where they need to be, they said, with just under half of students reading on grade level, they’ve come a long way: Reading scores are 4 percentage points higher than in 2012, and 13 points higher than six years ago. (Scores in math, science, and composition rose, also.)

Reflecting on the scores in an interview over the summer, Ms. Henderson said they show that the district’s investments in curricular materials, professional development, and good teachers and principals are starting to pay off.

“This isn’t episodic success, it’s systemic success,” the chancellor said. “We saw growth at every single grade, in every ward, all the subgroups. That tells me that it’s not just dependent on the quality of the principal in one particular school. This work is landing, and landing consistently across the board.”

The district leadership team is just beginning to mine the granular messages in the score data. In the coming weeks, officials will examine the literacy results strand by strand, said Brian Pick, who oversees curriculum for the district.

But a few big themes have already emerged and are shaping the school district’s approach to 2013-14. Sixth grade literacy was a weak spot citywide and will be drawing special attention as coaching and professional development moves forward. English-language learners and African-American boys, too, are not progressing well enough.

This year, the district hopes to spread and deepen its use of the response-to-intervention process—a screening strategy to see what students need—and put the right supports into place, Mr. Pick said.

“What this comes down to is building school teams’ abilities to look at individual kids and work with them and their families,” he said.

“It’s not rocket science; it’s having conversations as a team about what you are going to do to serve these kids.”

**Stories Within the Numbers**

The team at Stuart-Hobson, one of the higher-achieving middle schools in the 45,000-student system, found no shortage of reasons to celebrate as it sat down to peruse its test scores.

Schoolwide, the proficiency rate on the DC CAS in English/language arts rose from 59 percent in 2012 to 64 percent, with even brisker growth among black students and those from low-income families. Stuart-Hobson’s performance put it head and shoulders above its school district, whose K-12 reading proficiency rose from 43 percent to 47 percent.

There were things to celebrate that don’t show up on the year-end test, too. Suspensions, for instance, had dropped significantly, thanks apparently to a special focus on behavior issues by a new dean of students and a new assistant principal.

But some of the data points brought grievances. Proficiency rates declined for special education students. The 6th grade didn’t fare well in English/language arts. When that group of students finished 5th grade, 60 percent were proficient on the DC CAS; by the end of 6th grade, that proportion had dropped to 51 percent. A new teacher is now setting up shop in the 6th grade classroom at Stuart Hobson.

The 7th graders fared far better: Leaving 6th grade in 2012, 49 percent were reading on or above grade level. After a school year with teacher Kip Plaisted, 72 percent hit that mark. “He knocked it out of the park,” Principal Clemens said as she reviewed those figures.

Ms. McNair-Lee took her 8th graders from 60 percent proficiency to 69 percent.

With those kinds of numbers in hand, the Stuart-Hobson staff set about filling in a grid listing progress and challenges. In the coming weeks, that would morph into a school plan with goals for the new year that would be reviewed, in one-on-one meetings, by the regional superintendent and the school chancellor.

But now, the staff members detailed, subgroup by subgroup, student by student, the “glows and grows” they found in the data.

Analyzing one 6th grade group’s work, for instance, they wrote on the “glows” side of the grid that six of 41 moved to proficient, and one special education student remained at proficient. On the “grows” side—a reference to areas needing improvement—the teachers and administrators noted that five dropped from proficient to basic, 14 remained at basic, and two dropped to below basic.

In this way, tiny detail by tiny detail, they put their work ahead now, to her new role at Stuart-Hobson: overseeing a schoolwide enrichment program.

All around her, colleagues are parsing the minutiae of the test scores. At one table, the new 6th grade teacher, Matt Foster, works on a laptop with Stuart-Hobson’s reading-intervention specialist, Beth Dewhurst, using the forensic-data tool to locate test scores for all incoming 6th graders.

Getting a glimpse this way of his incoming students, Mr. Foster takes careful notes. It would be even better to have the data earlier, though, Ms. Dewhurst says.

“Next year, we’re going to make nice in the spring with all the principals [in feeder schools that] send us students. That way, well have this pipeline open earlier,” she said.

Across the room, Mr. Plaisted, the 7th grade English/language arts teacher, and Christopher Purdy, a special education teacher, are using the test-score data to brief Monica Green, who’s assumed Ms. McNair-Lee’s perch in 8th grade. It’s a passing of the torch.
Scrolling through the display on his laptop screen, Mr. Plaisted points one by one to students, discussing their scores, as well as their areas of academic strength and weakness, their home lives, personality quirks, and other information that will help Ms. Green anticipate their needs.

One student “is a sweet kid, and he tries,” Mr. Plaisted tells her. “We could tap into that.” Another student’s scores don’t reflect his reading ability. “I think he can read well,” Mr. Plaisted tells Ms. Green, “but when it comes down to the questions on the test, it throws him. It’s the test-taking.”

For other students, it’s the reverse: “This guy scored advanced only because he had a lucky day,” Mr. Plaisted says. And, “This kid never understood ‘main idea’ to save his life. He’s always focused on minute details.”

In this way, the teachers analyze every student at each level of performance, discussing the prospects for moving them further along.

A pivotal difference in their discussion flows from an important change this year to the district’s accountability system. Instead of being credited only for students who reach proficiency, schools now receive credit for each upward move a student makes on the performance scale.

The four levels on the test—below basic, basic, proficient, and advanced—are now divided into seven. That allows a school to earn 20 points on the district’s accountability index when a child moves up even within one of those bands, from “low below basic” to “high below basic,” for instance.

The change is meant to encourage teachers to focus not just on the “bubble kids” poised to move into proficiency, but on all students regardless of their place on the performance spectrum.

The August data sessions flowed into a schoolwide improvement plan for Stuart-Hobson that includes the goal of reaching 70 percent proficiency in reading this school year. Better coordinating the work of the three English/language arts teachers and working with teachers across the curriculum on techniques to help students master complex text will be key strategies in reaching that goal.

A New Focus

The District of Columbia system’s focus on the common standards in reading now moves into year 3, but layered on top is a push into the writing standards. That new priority was front and center in a late-August professional-development day.

Spread across classrooms on two floors of a high school, secondary-level teachers hunker down with instructional coaches to work on sentence composition.

Stuart-Hobson’s Matt Foster is here, with other 6th grade teachers, in a session co-led by Sarah Hawley, Stuart-Hobson’s assigned instructional coach. Mr. Plaisted joins 7th grade teachers across the hall, and Ms. Green does likewise with her 8th grade group.

Echoing the session leaders in the other rooms, Ms. Hawley guides her 6th grade teachers in an exercise about subordinating conjunctions. They’re learning how to work this kind of instruction into a class study of a text, instead of teaching it in isolation. They’re exploring how to “scaffold” the ideas, so all students can grasp them.

In the coming months, other professional-development sessions will focus on composing sentences and building paragraphs. Working with their own coaches at their schools this year, teachers will bring samples of their students’ work to analyze and to inform their instructional plans.

As Ms. Hawley begins shaping this year’s coaching plans for the teachers, she factors in a complex blend of teachers’ and students’ needs, test-score data, last year’s emphasis on text complexity and close reading, and this year’s move into the writing standards.

Two of the three English/language arts teachers are new to Stuart-Hobson this year, so new working relationships must be formed. There is much to do, and already the big clocks in each classroom serve as a constant reminder.

Other things have changed at Stuart-Hobson, too. Katie Franklin, who oversaw English/language arts as one of two assistant principals, is now in a district program that prepares her to become a principal. Inheriting her duties is Katherine Turner, an outgoing and energetic import from a nearby charter school.

Ms. McNair-Lee has given up her classroom to direct a new schoolwide enrichment program that allows students to study through the lens of something they’re interested in.

Even though Ms. McNair-Lee was ready for something new, leaving her classroom was still causing pangs in her belly as the new year began. She’ll keep her hand in teaching, though, working with students on projects.

She’ll think about last year’s students, the ones she greeted every morning with her standard line: “Good morning, scholars.” As the summer heat wanes, they’re learning their way around the unfamiliar campuses of high schools across the city.

One of those students—one of those she worries about most—is starting out in his new high school with an uncertain hold on important skills. And as Mikel disappears into those wide, crowded hallways, where no teacher knows his name yet, his former teacher wonders how he’s doing, and she crosses her fingers.

This story is part of a four part series, entitled “Common Core: A Steep Climb.” To view the entire story series, please visit: http://www.edweek.org/ew/ collections/common-core-a-steep-climb/index.html

Challenges Seen in Testing Special Ed. Pupils on Common Core

By Nirvi Shah

Washington

The path to devising assessments for students with disabilities that measure how adept they are at mastering the Common Core State Standards seems to be filled with hurdles to overcome before students face those assessments in the 2014-15 school year.

States are in the early stages of implementing the common standards, adopted by all but four states. Two consortia of states have been awarded contracts to design exams for most students—including some with disabilities—who will take the tests, which will be computer-based or computer-adaptive. Another two groups are designing exams based on the standards for the 1 percent of students with the most severe cognitive disabilities. All four groups are in various stages of test development.

One of the obstacles facing students with disabilities who will take the exams has less to do with the tests than with instruction, said Stephen N. Elliott, an education professor at Arizona State University, in Tempe. Mr. Elliott spoke May 22 at a U.S. Department of Education meeting addressing the challenges that remain in preparing new tests that all students are scheduled to take in 2014.

In his research, Mr. Elliott has found that the most time any state was able
to spend on teaching the current standards was 81 percent of the time students were in school, and special education teachers covered even less of the content and standards.

“We get that test score, and we make that big inference that kids have been taught this,” Mr. Elliott told the gathering of special education and testing experts, including members of the consortia that are designing common-core assessments and alternate assessments for students with significant cognitive disabilities.

“Many students with disabilities need 30 to 40 more days of class time to get an equitable opportunity to learn.”

And that disparity may only grow as the demanding common standards, in English/language arts and mathematics, are put in place.

“It’s not that we can’t improve assessments, [but] that can serve as a distraction from the critical need to improve instruction,” said Louis Danielson, a managing director who focuses on special education policy and evaluation at the American Institutes for Research, a Washington-based organization.

Progress and Problems

The major hurdle of increased, improved instruction aside, the technical and content issues posed by the exams are numerous, experts at the Education Department forum said.

Students with disabilities have become a bigger part of state accountability systems, albeit gradually, during the past 20 years, so that now even students with the most significant cognitive disabilities are included in state testing programs.

One fundamental advantage to designing tests with students with disabilities in mind from the beginning is that, for the most part, the tests won’t have to be adapted to work with those students after the fact, disability education experts have said. A need for such retrofitting is common with current state assessments.

But there’s still a long way to go, Mr. Danielson said.

One big issue lies with computer-adaptive tests, which pull from a bank of test questions with a wide range of difficulty. The computer adjusts the difficulty of the questions it poses based on a student’s performance on previous questions. One problem with that approach is that some students may shut down if they miss the first question, Mr. Danielson said. Then there’s the risk that the computer will throw a student a question that’s below his or her grade level because of a series of incorrect answers that leads the computer to those questions, a possibility that concerns special education advocates.

Yet another issue is that states using exams developed for most students by one of the two consortia working on those tests will have to agree on a common set of acceptable test accommodations—adjustments made, in other words, to help students with disabilities access the test content as easily as classmates without disabilities.

Read-Aloud Debate

Common accommodations include giving students additional time to take an exam, giving them a separate testing area, limiting questions to appearing one at a time, and adjusting the size of the typeface of the test. But one accommodation over which there is disagreement is whether, or how much, students should have test instructions or test content read aloud to them.

“I feel when reading reaches a point where it’s about comprehension and they still have trouble decoding it, it becomes a test of decoding,” Mr. Danielson said. “In earlier grades where decoding is being tested, it makes sense not to read aloud.”

Students may waste a lot of their time when they hit unfamiliar proper nouns, reducing their fluency and comprehension, he said research shows. And students using digital text in class where read-aloud features are common may be stumped on tests where those features aren’t allowed.

Alexa Posny, the Education Department’s assistant secretary for special education and rehabilitative services, said one proposed accommodation has already been a subject of concern: sign-language avatars. Students could invoke an avatar—a computer-generated version of a sign-language interpreter—to convert parts of the test into sign language if that accommodation was deemed necessary. (See Education Week, June 8, 2011.)

However, if students are accustomed to live sign-language interpreters in class, the avatars could be startling and awkward to use.

“Students shouldn’t suddenly encounter an accommodation they haven’t used in the classroom,” said Sheryl Lazarus, a senior research assistant at the National Center on Educational Outcomes, in Minneapolis, which is leading one of two groups of states in designing alternative assessments for students with severe cognitive disabilities.

Those alternatives are a separate set of exams, now often referred to as “1 percent” tests. That’s because the federal No Child Left Behind Act allows the scores of no more than 1 percent of students who take such exams to count in school accountability measures, in part to discourage schools from giving the alternatives to too many students.

In some cases now, students with severe cognitive disabilities don’t take exams that resemble those of peers without disabilities. Instead, a collection of their work is put together to demonstrate their skills.

Mr. Danielson argues that the time has come to put an end to those portfolio-style assessments—a step that would not be universally popular. He said there are too many questions about how reflective of students’ ability such collections are and how heavily they are influenced by teachers, who may be evaluated on their contents. Overly high proficiency rates may signal that expectations for students with disabilities are too low, he said.

“It’s too tempting for teachers to help a lot with their [students’] work,” Mr. Danielson said.

It’s not that we can’t improve assessments, [but] that can serve as a distraction from the critical need to improve instruction.”

LOUIS DANIELSON
Managing Director, American Institutes for Research
Great Books Programs Prepare Your K–8 Students for the Upcoming PARCC Assessments!

Meeting Your Needs

The Great Books Foundation provides strong, inquiry-based English language arts programs for grades K–12 that develop critical literacy skills in reading, writing, listening/speaking, and language. We believe these skills will enable students to perform better on the PARCC assessments that will be used by most states beginning in the 2014–2015 school year.

To find out how Great Books programs prepare your students for the PARCC assessments, visit www.greatbooks.org/parcc.

Great Books Programs

We support teachers and students with a systematic instructional approach that reflects the rigor of the Common Core State Standards and PARCC assessments. In our programs, students grapple with close, analytic reading of complex texts and construct increasingly sophisticated responses in speaking and writing.

The PARCC Model Content Frameworks, a guide for bridging instruction and assessment, identifies the following key elements and critical emphases in both the Common Core State Standards and PARCC assessments:

**Read complex texts**
- Reading and comprehension of a range of complex texts from many domains
- Critical reading comprehension, assessed in context of reading passages.
- Close, analytic reading and comparing and synthesizing ideas across texts
- Reading of short passages and longer passages
- Engagement with complex, content-rich, and literary texts to build knowledge

**Write to texts**
- Analytic writing with evidence to support an explanation, summary, claim, or comparison
- Narrative writing of stories and descriptions based on reading
- Careful, close reading and clear, coherent writing for both analytic and narrative writing

**Research**
- Conducting and reporting on research to answer questions or solve problems
- Gathering, analyzing, and evaluating resources
- Reporting on information and ideas investigated
Read Complex Text: Literature

<table>
<thead>
<tr>
<th>Characteristics of Great Books Programs for Grades 3–5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Great Books programs offer complex, diverse, thematically rich literature from renowned authors that sparks rigorous thinking and discussion</td>
</tr>
<tr>
<td>2. Increasingly challenging selections progress in reading level, conceptual complexity, and length throughout the series</td>
</tr>
<tr>
<td>3. Shared Inquiry activities with a recognizable structure for deep comprehension: two readings, questioning, vocabulary, note-taking activities, Shared Inquiry discussion, writing, and related reading</td>
</tr>
</tbody>
</table>
| 4. Shared Inquiry offers a systematic approach to solve central problems of meaning in text using three elements of critical thinking:  
  • IDEA—generate/clarify ideas about meaning  
  • EVIDENCE—support ideas with text evidence  
  • RESPONSE—consider others’ ideas and modify interpretation |
| 5. Shared Inquiry discussion, the cornerstone activity for collaborative analysis and discovery of meaning in complex literature with consideration of ideas and perspectives of others |
| 6. Directed note-taking activities that trace the development of important concepts and ideas across the text, including comparison and contrast of characters to formulate interpretations and judgments |
| 7. Follow-up questioning prompts tailored to student responses to deepen understanding of the text |
| 8. Multiple close readings of every text to uncover layers of meaning, raise thought-provoking interpretive questions, and build interpretations |
| 9. An abundance of tools and activities for gathering evidence for use in writing, such as evidence organizers and directed note-taking |
| 10. Challenging interpretive questions for each selection allow students to explore, substantiate, and formulate multiple hypotheses about a piece of literature |
| 11. Interpretive activities to determine two or more main ideas in a text using key details |
| 12. Cross-text activities in Story-to-Story Connections to compare and contrast multiple texts |
| 13. A clear focus on the impact of specific words and details at the sentence and passage level to analyze how the author creates character, setting, mood, and tone |
| 14. Multiple close readings using directed notes for examination and analysis of how the subtleties of language affect meaning |
| 15. Interpretive questioning prompts to analyze text details, determine importance of character actions and story events, and build a coherent argument |
| 16. Fluency practice using syntax, meaning, and punctuation with attention to expression and phrasing, supported by professionally recorded CDs as models |
| 17. Ongoing informal assessments include clear outcomes with a spectrum of student behaviors and reflection tools to measure progress and set goals for reading and discussion |
| 18. More formal assessments include story-specific comprehension tests and a critical thinking rubric |
### Write to Text: All Writing

**Characteristics of Great Books Programs for Grades 6–8**

1. Great Books writing activities focus on the use of credible text evidence to produce clear, coherent analytic and narrative writing appropriate to task, purpose, and audience.

2. Writing embedded in the Shared Inquiry approach includes responses to prereading prompts, note taking, and question generating.

3. Activities to help students build answers using evidence organizers and drafting guides.

4. Writing opportunities to extend or consolidate interpretations about text after peer and teacher response.

5. Variety of writing forms including paragraphs, essays, letters, journal entries, poems, and stories.

6. Activities for students to write imaginative exercises with characters in alternate settings, scripts for dramatic scenes from texts, and dialogue between characters from different texts.

7. Multi-genre products with visual arts and writing combinations to create comic strips, and theatrical and graphic representations of characters with explanations, and setting maps with navigational instructions.

8. Multiple opportunities and ways to write about characters to enhance understanding of texts.

9. Apprentice method to learn how to write in different genres with emphasis on imitating a style using a writer’s own experience.

10. Opportunities to foster understanding of audience through the use of portfolios, rubrics, and peer and teacher feedback.

11. Variety of structured writing experiences:
   - Expository—evaluative, analytic, and interpretive essays
   - Creative—short stories, scripts, guide books, and reimagined events in alternative settings

12. Extensive writing support with process guides, evidence organizers, peer response forms, and rubrics.

13. Consistent, structured writing practice to text-specific prompts for comprehension and critical thinking using graphic organizers.

14. Complex, authentic fiction and nonfiction reading selections as models for writing in a variety of genres.

### Write to Text: Analytic and Informative/Explanatory Writing

**Characteristics of Great Books Programs for Grades 6–8**

1. Written response to rigorous text-dependent prompts to frame a valid argument or to inform the reader of what has been learned.

2. Writing interpretive essays to make arguments with a more sophisticated, selective use of evidence and a coherent organizational structure.

3. Writing evaluative essays with a clear thesis; coherent, developed support; and a conclusion.

4. Focus on writing skills that are the building blocks of analytic writing: comparison and contrast, cause and effect, and definition and classification.

5. Writing reflections about text-specific features, such as the author’s use of language to create mood, tone, point of view, character development, and movement of events.
Great Books and the PARCC Assessments, grades K–2 excerpt

Although PARCC assessments have not yet been developed for grades K–2, Great Books reading foundation activities will prepare students for PARCC assessments before they enter third grade.

**Listening and Speaking**

<table>
<thead>
<tr>
<th>Characteristics of Great Books Programs for Grades K–2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Extensive collaborative critical thinking and discussion opportunities about text, through Shared Inquiry discussion and interpretive activities with peers and adults in partnered, group, and teacher-led settings</td>
</tr>
<tr>
<td>2. Shared Inquiry discussion, the cornerstone activity for collaborative analysis and the discovery of meaning in complex literature with consideration of the ideas and perspectives of others</td>
</tr>
<tr>
<td>3. Discussion guidelines and mini-lessons that foster collaborative discussion behaviors and respectful speaking/listening skills</td>
</tr>
<tr>
<td>4. Opportunities for adults to model specific behaviors that encourage respectful discussion of varied perspectives through questioning, thinking, speaking, and listening skills</td>
</tr>
<tr>
<td>5. Interactive oral communication and interpersonal skill development opportunities at every stage of the units—from asking questions and expressing a variety of ideas with key details to building interpretations and formulating conclusions through Shared Inquiry discussions</td>
</tr>
<tr>
<td>6. Shared Inquiry discussions that rely on collaboration and communication to solve problems of meaning</td>
</tr>
<tr>
<td>7. Shared Inquiry discussion and activities that provide extensive and consistent practice of critical listening skills. Students will:</td>
</tr>
<tr>
<td>• Ask and answer questions about a text that is read aloud or a speaker’s comments</td>
</tr>
<tr>
<td>• Request clarification, gather information, and clarify comprehension</td>
</tr>
<tr>
<td>• Modify thinking after hearing a variety of perspectives and arguments</td>
</tr>
<tr>
<td>8. Text analysis activities and question prompts that teach close, critical listening as rich passages are reread and discussed</td>
</tr>
<tr>
<td>9. Shared Inquiry discussion and activities that provide extensive and consistent practice of critical speaking skills. Students learn to:</td>
</tr>
<tr>
<td>• Express their own ideas clearly</td>
</tr>
<tr>
<td>• Ask and answer questions</td>
</tr>
<tr>
<td>• Provide explanations</td>
</tr>
<tr>
<td>• Build on others’ ideas</td>
</tr>
<tr>
<td>• Agree and disagree respectfully</td>
</tr>
<tr>
<td>10. Carefully crafted questioning prompts, follow-up questions, and activities that encourage further clarification of ideas and more detailed, relevant text support with fuller explanations</td>
</tr>
<tr>
<td>11. Shared Inquiry discussions that prompt students to give reasons for their answers by referring to the text</td>
</tr>
<tr>
<td>12. Fluency practice of key passages to develop expression and confidence with attention to phrasing, syntax, meaning, and punctuation, supported by professionally recorded CDs as models</td>
</tr>
<tr>
<td>13. Interactive text-related activities and projects using a variety of visual, verbal, kinesthetic, and written modalities for collaboration, interpretation, self-expression, and presentation</td>
</tr>
<tr>
<td>14. Opportunities at the end of each unit to present and listen to findings about text-related topics from cross curricular projects</td>
</tr>
<tr>
<td>15. Student and teacher reflection activities and tools to personally assess discussion and critical thinking skills, identify strengths and challenges, and set goals</td>
</tr>
<tr>
<td>16. Parent component with guidelines and text-specific questions that provides additional comprehension and critical listening and speaking practice</td>
</tr>
</tbody>
</table>
Putting the brakes on common-core stakes could stop the clock on progress

By Alice Johnson Cain

President Randi Weingarten of the American Federation of Teachers recently proposed a moratorium on “high stakes” related to the Common Core State Standards, positing that the standards will either lead to a revolution in teaching and learning or end up in the dustbin of abandoned reforms. She is right that the potential for a transformation—one of particular benefit to children who are not well served by our current education system—is real. But a moratorium would be a mistake.

The common core is not just another reform; it is truly a revolutionary development. But it is also a package deal in which next-generation assessments will inform and improve instruction in ways that make far more sense to teachers than the current “bubble tests” that are often disconnected from what they teach and what their students need.

The common core defines critical, real-world understandings that students need for success in college and career, broken down by grade level. Until the development of the common standards in English/language arts and mathematics, teachers have been expected to cover a surfeit of material. The common core reduces this load, but in return demands that teaching be pursued with increased rigor, depth, and relevance. This call for excellence has been met with increased rigor, depth, and relevance.

The teachers we work with in my organization, Teach Plus, are eager for the new assessments that will accompany the common core, for the first time in history, teachers across the United States are united in a way that opens the door for the use of massive online open courses, or MOOCs, as professional-development tools.

This means some of the nation’s best teachers can lead online professional development for thousands of their peers for a particular grade level on a particular standard. As Latisha Coleman, a 2nd grade teacher in Washington’s Inspired Teaching Demonstration Public Charter School, put it during a recent meeting convened by Teach Plus: “Common core has created a common language for teachers. It has broadened our ability to share best practices across schools and across systems.” This could have a revolutionary impact on the teaching profession.

Of course, enthusiasm for the new standards doesn’t mitigate all concerns. Like Weingarten, many teachers are rightly worried that, in an effort to speed implementation, districts will fail to provide the level and quality of support that teachers need. But delaying something is not the same thing as improving it. Instead, districts should find ways to empower and invest in teacher-leaders, who would make ideal trainers for their colleagues. This would enable teachers to learn about the standards from in-the-classroom experts they trust, instead of outsiders. Hybrid roles for teacher-leaders, so that they can teach part time while also coaching colleagues, is one way we can invest in teachers as experts to make the common-core rollout a success without wasting time.

The teachers we work with in my organization are enthusiastic about the standards because they offer opportunities for cross-country collaboration on best practices and for high expectations for students’ critical-thinking skills—the learning that matters most.

Likewise, many teachers are eager for the new assessments that will accompany the common core because they squarely address the biggest complaint against No Child Left Behind-era tests—that the focus on multiple-choice standardized tests has made a test score, not authentic learning, the goal of education.

The new, computer-based assessments aligned to the common core are as far from fill-in-the-bubble tests as this nation has ever seen. Many of the questions on them are open-ended, and the assessments share an emphasis on showing an understanding of ideas, not just identifying an answer. Best of all, they are designed to measure deep knowledge and complex, higher-order thinking and problem-solving skills. No doubt the new assessments won’t entirely rid our schools of the occasional lousy (and sometimes downright irresponsible) use of testing; policy is never fool-proof. But that alone is not a reason to “put the brakes on stakes” (as moratorium proponents put it) just at the moment when new assessments are finally going to be far better measures of student learning than NCLB-era tests ever were.

Teaching to the improved standards without the aligned assessments—or with assessments that don’t “count”—would lay the groundwork for shortsighted decisions to skip the assessments altogether or the continuous delay of that magic day when they will count, in an endless quest to attain perfection. The bumpy road that led to recent steps forward in developing evaluation systems that are based in part on student learning gains could be erased.

When asked about the proposed moratorium, New York state Commissioner of Education John King quoted Martin Luther King Jr.’s famous “Letter from a Birmingham Jail”: “Frankly, I have yet to engage in a ... campaign that was ‘well timed’ in the view of those who have not suffered unduly. ... For years now I have heard the word ‘Wait!’ It rings in the ear ... with piercing familiarity. This ‘Wait’ has almost always meant ‘Never.’” Children who are still waiting for an education that prepares them for success in college and careers cannot afford the risk of continued delay.

But Randi Weingarten is right that supporting teachers every step along the standards path is essential. The federal government sends states $2.5 billion each year for professional development for teachers and gives great latitude to states and districts
Note to Congress: Fewer, Better Tests Can Boost Student Achievement

By Marc Tucker, Linda Darling-Hammond, and John Jackson

Both Democrats and Republicans have submitted proposals to reauthorize the Elementary and Secondary Education Act, the federal law governing K-12 education that has not been revamped since the much-maligned No Child Left Behind Act—the latest iteration of the ESEA—was signed into law in 2002. Among the few things both parties agree on is continuing to require grade-by-grade testing and a new requirement that would focus the draconian consequences that once threatened all schools on the lowest-achieving schools exclusively; in other words, those schools that primarily serve low-income and new immigrant students. This is a big mistake.

Taken together, along with our continued failure to address equity in school resources, these two provisions would virtually guarantee that the overall performance of our students will never equal that of our toughest international competitors and would further widen the gap between the top performers and our disadvantaged students.

Here’s why: Americans are addicted to multiple-choice, computer-scored tests, mainly because they are cheap and easy to score. However, these tests drive a rote curriculum that will not produce the skills students need to get and keep good jobs in the 21st century—writing and speaking well, using advanced mathematics, analyzing complex problems, and finding and synthesizing information from many sources for creative problem-solving.

The countries that outperform the United States on international exams spend more than we do to measure and encourage these skills with essay tests and teacher-scored projects. And they can afford to do this because they test much less frequently than we do, typically only two or three times during a student’s entire school career.

Although the new Common Core State Standards outline ambitious skills, states worried about the costs of grade-by-grade testing are putting pressure on the state consortia developing new tests—the Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Careers, or PARCC—to lower their costs, even though this would eliminate the kinds of tasks that can measure the most challenging standards. Commercial testing companies are likely to follow suit, with the result that our assessments will change far too little to match those of high-achieving nations.

We get what we measure. This is truer now than ever in the United States because of new federally required accountability systems that tie teachers’ employment and some schools’ existence to student test performance. By continuing these accountability requirements and the requirements for annual high-stakes testing that effectively force states to use cheap tests, Congress will virtually guarantee that our teachers have strong incentives to teach a curriculum that leaves out the complex skills and knowledge most needed by our students in the modern global economy.

Moreover, by refocusing sanctions on schools serving the most disadvantaged students, Congress is making sure that the students who most need a top-flight education to create a bridge to the jobs of the future will, in fact, receive the most limited curriculum.

If, at this critical moment, the United States chooses to measure 20th-century skills when our top competitors are measuring 21st-century skills, we will effectively be putting a cap on what the American education system will be able to deliver, sealing the fate of our most disadvantaged students and likely of the entire country.

There is a simple fix. Absent major federal
investments in much better assessments, Congress should require that states ensure external testing for a single grade at each of three school levels—elementary, middle, and high school—while continuing to report scores separately for vulnerable groups. Without spending any more than they do now, states could employ higher-quality assessments that encourage more productive teaching while reducing the testing burden on students and teachers.

While some will argue that reducing the number of grades tested would reduce accountability and weaken incentives to improve performance, there is no evidence here or around the globe that more testing produces higher achievement.

The record of the top-performing countries is clear—high standards aligned to first-rate curricula and assessments are a necessary step to enable high and equitable achievement. To match the academic performance of these countries, we will have to adopt the other key contributors to their success: fairer systems of school finance, more support for early-childhood education, more coherent and powerful instructional systems, and a whole galaxy of policies aimed at recruiting and training top-notch teachers and fairly distributing those teachers among schools and students. If we are to succeed, we will need to create tests worth teaching to and implement the common-core standards along with common-core supports.

Marc Tucker is the president of the Washington-based National Center on Education and the Economy. Linda Darling-Hammond is the Charles Ducommun professor of education at Stanford University and the faculty director of the Stanford Center for Opportunity Policy in Education. She is a member of the Technical Advisory Committee of the Smarter Balanced Assessment Consortium. This Commentary represents Ms. Darling-Hammond’s personal views and not the views of the consortium. John Jackson is the president of the Schott Foundation for Public Education, based in Cambridge, Mass.
Get the information and perspective you need on the education issues you care about most with Education Week Spotlights

The Achievement Gap • Assessment • Autism • Bullying • Charter School Leadership • Classroom Management • Common Standards • Common-Core Assessment and Development • Differentiated Instruction • Dropout Prevention • E-Learning • ELL Assessment and Teaching • ELLs in the Classroom • Flu and Schools • Getting The Most From Your IT Budget • Gifted Education • Homework • Implementing Common Standards • Inclusion and Assistive Technology • Literacy and the Common-Core • Math and the Common-Core • Middle and High School Literacy • Motivation • No Child Left Behind • Pay for Performance • Principals • Parental Involvement • Race to the Top • Reading Instruction • Reinventing Professional Development • Response to Intervention • School Uniforms and Dress Codes • Supporting Instruction in the Common Core • Teacher Evaluation • Teacher Tips for the New Year • Technology in the Classroom • Tips for New Teachers

VIEW THE COMPLETE COLLECTION OF EDUCATION WEEK SPOTLIGHTS

www.edweek.org/go/spotlights