



HARVARD EDUCATION LETTER

I N S I D E

interview

**Ronald Ferguson on
Lifestyle, Classroom Culture,
and Achievement Gaps 4**

point of view

**Richard F. Elmore on
Lost Instructional Time 8**

(In)formative Assessments

New tests and activities can help teachers guide student learning

by Robert Rothman

Although many teachers in the No Child Left Behind (NCLB) era complain that students take too many tests, teachers at the John D. Philbrick Elementary School in Boston eagerly signed on last year to give students six *more* tests a year. The tests, known as Formative Assessments of Student Thinking in Reading, or FAST-R, are short multiple-choice quizzes that probe key reading skills. The tests are designed so that teachers can make adjustments to their instruction based on students' answers.

With FAST-R "we get concrete, helpful information on students very quickly," says Steve Zrike, Philbrick's principal.

Now used in 46 schools in Boston, FAST-R is part of a rapidly growing nationwide effort to implement so-called formative assessments—tests that can inform instruction through timely feedback. (By contrast, end-of-term tests and standards-based accountability tests are called summative assessments because they provide a summary of what students have learned.) Interest in formative assessment is fueled by the growing pressure to raise student achievement. Because the state tests on which schools will be judged under NCLB are typically administered at the end of the academic year, educators like Zrike are eager for information that can help them predict whether students are on track toward meeting proficiency goals and then intervene appropriately.

A strong body of evidence indicates that formative assessment, done properly, can generate dramatic improvements in teaching and learning. But some experts warn that many of the instruments marketed as formative assessments are in effect summative tests in disguise. At best such tests provide little useful information to class-

room teachers; at worst they can narrow the curriculum and exacerbate the negative effects of teaching to the test, says Lorrie A. Shepard, dean of the School of Education at the University of Colorado at Boulder.

"If all the test produces is a predictive score, or tells you which students to be anxious about, it's a waste of money," she says. True formative assessments, she says, tell teachers "what it is the students aren't understanding."

"If pilots navigated the way [most] teachers teach, they would leave London, head west, and at the end of eight hours, ask, 'Is this New York?'"

Tasting the Soup

With the proliferation of so many instruments, it's not surprising that many educators find the distinction between formative and summative assessment confusing (see "Three Types of Assessment," p. 2). Paul Black, coauthor of a landmark 1998 study on the topic, once described the difference by saying that formative assessment is when the chef tastes the soup; summative assessment is when the customer tastes the soup. As his remark implies, the effectiveness of the process depends not only on the data sampled, but the timeliness of the feedback and how the "chefs"—not only teachers but the students themselves—use it.

In the 1998 study, Black and Dylan Wiliam, both of Kings College, London, examined some 250 studies from around the world and found that the use of formative assessment techniques produced significant accelerations in learning. Students in classes using this approach gained a year's worth of learning in six or seven months. The method appeared particularly effective for low-performing pupils. As a result, formative assessment was found to narrow achievement gaps while it raised achievement overall.

"We know of no other way of raising [achievement] standards," the authors conclude, "for which such a strong



**Harvard
Education Letter**

EDITOR
Suzanne Clayton

EDITOR
Arlene Chauncey

ASSISTANT EDITOR
Deborah Walser

PRODUCTION MANAGER
Deborah Riggs

EDITOR
David F. Elmore

EDITORIAL ADVISORY BOARD
Bernice C. Boies, Lecturer, HGSE;
Lara A. Cooper, Asst. Superintendent,
Evanston Township (Ill.) High School;
Linda Darling-Hammond, Professor,
Stanford University; Sally Kagan,
Vice President for Programs and
Partnerships in Education, Boston
College; Susan Foreman Johnson,
Professor, HGSE; Robert Kagan,
Professor, HGSE; Peggy Kemp,
Head of School, Fenwick High School,
Boston; Marya Benson, Director,
Teacher Education, Brandeis University;
Deborah Carter, Co-Principal, Mission Hill
School, Boston; John Merrow, President,
The Merrow Report; Jerome Murphy,
Professor, HGSE; Arthur Loewenthal,
Publishing Consultant; Bernice Snow,
Professor, HGSE; Richard Sugarman,
Teacher, Runkle School, Brookline,
Mass.; Ariadne Tsamis, Director of
Program and Source Development,
John F. Kennedy Library Foundation.

Harvard Education Letter (ISSN 0895-3716) is published monthly by the Harvard Graduate School of Education, 8 Story St., Cambridge, MA 02138-3752. Second-class postage paid at Boston, MA, and additional mailing offices. Postmaster: Send address changes to Harvard Education Letter, 8 Story St., Cambridge, MA 02138-3752.

Reprinted articles in the Harvard Education Letter represent the views of the authors.

© 2006 by the President and Fellows of Harvard College. Published as a non-profit service. All rights reserved. Special permission is required to reproduce in any manner, in whole or in part, the material herein contained.

HOW TO SUBSCRIBE
Send \$38 for individuals, \$49 for institutions (\$52 for Canada/mexico, \$63 other foreign, in U.S. funds only). Subscription prices subject to change without notice. Single copies, \$7.00. Back issues and bulk subscriptions available at special reduced rates.

Address all correspondence to Harvard Education Letter, 8 Story St., Cambridge, MA 02138-3752; phone 617-355-3432 in Massachusetts, 617-552-0763 outside Mass.; fax 617-496-3584; email: editor@edletter.org; web: www.edletter.org.

prima facie case can be made on the basis of evidence of such large learning gains."

While skilled teachers may be adept at checking for comprehension, identifying misunderstandings, and adjusting instruction accordingly, the authors cited ample research showing the pervasiveness of ineffective or counterproductive assessment techniques in the classroom. "If pilots navigated the way [most] teachers teach," says William, now deputy director of the Institute of Education in London and former senior research director for the Educational Testing Service (ETS), "they would leave London, head west, and at the end of eight hours, ask, 'Is this New York?'"

Even teachers who check for students' understanding at the end of every lesson seldom get enough information to guide instruction, he adds. "They make up a question at the spur of the moment, ask the whole class, six kids raise their hands, and one answers," William says. "That's what I did as a teacher. But how dumb is that?"

"Teachers need better data to make instructional decisions," he adds.

The Uses of Benchmark Testing

William claims his research has been misinterpreted to suggest that *any* periodic assessment is an effective intervention. Jumping on the formative assessment bandwagon, test publishers have begun selling benchmark or early warning assessments linked to their end-of-year tests that indicate whether students are on track to pass. According to Tim Wiley, a senior analyst at Eduventures, LLC, these assessments represent the fastest-growing segment of the testing industry. Total spending on such instruments is approximately \$150 million.

Stuart Kahl, president and CEO of Measured Progress, a testing firm based in Dover, N.H., cautions that benchmark tests are designed to obtain information about groups rather than individuals and should not be confused with formative assessment. The tests usually include only a handful of items on each topic in order to survey knowledge across an entire unit. But while one or two items can provide information on the skills of an entire class or a school, they do not yield enough information about an individual student's understanding to

guide instructional decisions. For instance, if three quarters of the class miss both questions on multiplication, the teacher knows she needs to revisit this topic. But are these computational errors or conceptual problems? This kind of test is not likely to reveal the answer for any particular student.

Measured Progress produces a series of benchmark tests known as Progress Toward Standards. Kahl notes that these kinds of tests can be useful for interim program evaluation and for identifying patterns of performance. For example, if girls score better than boys across the board, that may spur schools to examine curriculum and instructional practices. But Kahl agrees with the University of Colorado's Shepard that the interim tests do not provide the type of information about individual student progress that appropriate formative assessments provide.

"Formative assessment is a range of activities at the classroom level [that] teachers use day in and day out to see if kids are getting it while they're teaching it," he says.

Formative Assessment Techniques

Districts and schools using formative assessments employ a variety of techniques. For instance, commercial companies are developing a variety of new assessment tools, ranging from handheld electronic "clickers" that allow students to register their responses to teachers' questions to instructional software programs that incorporate checks on student understanding. This helps teachers gauge the progress of the entire class, not just the students who raise their hands. Some six million students are using such programs at a cost of about \$120 million, according to Eduventures' Wiley.

Some curriculum programs also include formative assessment techniques to help teachers gauge student understanding while they are teaching, notes Shepard. For example, a mathematics program might ask students to multiply 3 by 4 in three different ways: by making sets of three, by calculating the area of a floor, and by counting by fours. A right answer on any of these indicates that a student grasps the concept of multiplication. Otherwise, Shepard says, "you don't know if the student doesn't understand the concept or multiplication facts."

State tests provide too little information, and too late, to help teachers in the classroom. By the time the results come back "kids are long gone, in another class."

Three Types of Assessment

	<i>Summative</i>	<i>Benchmark</i>	<i>Formative</i>
Key Question	Do you understand? (yes or no)	Is the class on track for proficiency?	What do you understand?
When Asked	End of unit/term/year	6-10 times per year	Ongoing
Timing of Results	After instruction ends	Slight delay	Immediate

At ETS, William developed a series of workshops called Keeping Learning on Track to help teachers develop formative assessment strategies and monitor their own progress. These workshops have been conducted in 28 districts, including Cleveland (see "Helping Students Assess Themselves"). The workshops revolve around five key strategies: sharing expectations for learning, effective questioning, providing meaningful feedback, student self-assessment, and peer assessment among students. In each district, teachers come together every month to discuss how they implemented the strategies and the results they produced.

Donna Snodgrass, executive director for standards, curriculum, and classroom assessment for the Cleveland public school system, says the effort is essential because state tests provide too little information, and too late, to help teachers in the classroom. By the time the results come back, she notes, "kids are long gone, in another class."

"This gives them an idea of where kids are and what they can do about it," she says.

And it appears to be working. After two years, students' mathematics scores in the ten participating Cleveland schools—among the lowest-performing in the city—rose four times faster than those in comparable schools. And while the program is aimed particularly at mathematics instruction, the program has also had an effect on reading achievement: Reading scores in the ten schools increased four to five percentage points over two years. Snodgrass says these kinds of results have encouraged teachers to keep trying new strategies. The district is also planning to expand the program to additional schools.

Insight into Student Mistakes

Other districts, like Boston, use more formal assessment instruments. FAST-R, which was developed by the Boston Plan for Excellence, uses ten multiple-choice questions to probe student comprehension of a particular reading passage. The test can be used as a benchmark assessment, since it is aligned to state end-of-year tests, but it is also designed to help improve reading instruction, says Lisa Lineweaver, a senior program officer at the Boston Plan. It focuses in depth on only two skills (finding evidence and making inferences) and helps teachers diagnose reading difficulties on the basis of wrong answers. Answers are categorized as correct; "out of place," meaning that the answer is "a near miss" based on a misreading of the text; or "out of bounds," meaning that it is not based on the text. Comments on particular wrong answers can help teachers see where students are having trouble. Are they associating one word with another? "Plugging in" a plausible but irrel-

Helping Students Assess Themselves

In their review of studies on formative assessment, Paul Black and Dylan William stress that self-assessment by students is "an essential component" of formative assessment. What students need, they argue, is a clear understanding of their learning goals, useful feedback on where they stand, and some sense of how they can close that gap.

"Surprisingly, and sadly," they write, "many pupils do not have such a picture, and appear to have become accustomed to receiving classroom teaching as arbitrary sequences of exercises with no overarching rationale."

Keeping Learning on Track, a series of formative assessment workshops William developed for Educational Testing Service (ETS), helps schools provide opportunities for students to gauge their own learning. In Cleveland, for example, teachers begin by *making explicit their daily learning goals and posting examples of high-quality work* that meets those goals. This helps students judge not only their own work but also that of their peers.

In place of letter grades, smiley faces, or check marks, the Cleveland teachers *provide substantive feedback* on students' work. They return each paper with "two stars and a wish": the stars indicate things the student did well, and the wish is for a substantive improvement.

Students are also encouraged to *seek out learning help from their peers*. In the course of a lesson, students put a green cup on their desk when they feel confident they understand the new material. The teacher tells students who still have questions to "ask three [students] before you ask me," according to Donna Snodgrass, executive director for standards, curriculum, and classroom assessment for the Cleveland public schools.

To *hold students accountable for their learning*, Cleveland teachers call on students randomly by drawing popsicle sticks with names on them. Snodgrass notes that ETS researchers used a similar technique when introducing the method with school principals. "Everybody put their paperwork away," she says. "It was a good object lesson." ■

evant answer? Misapplying information gleaned elsewhere in the text?

The results come back a few days after the test is administered. After the assessment, the Boston Plan provides coaches who work with teachers to help them adjust instruction appropriately, based on student responses. Teachers are also encouraged to conference with children to probe more deeply into students' level of background knowledge and reasoning processes. "There is no way of knowing for sure what a kid is thinking unless you know the kid," Lineweaver says. ■

Robert Rothman is a principal associate at the Annenberg Institute for School Reform at Brown University and the editor of *Voices in Urban Education*.

For Further Information



P. Black and D. William. "Assessment and Classroom Learning." *Assessment in Education* 5, no. 1 (1998): 7-74.

P. Black and D. William. "Inside the Black Box: Raising Standards through Classroom Assessment." *Phi Delta Kappan* 80, no. 2 (1998): 139-148. Available online at <http://www.pdkintl.org/kappan/kbla9810.htm>.

L. A. Shepard. "The Role of Classroom Assessment in Teaching and Learning." (CSE Technical Report 517.) Los Angeles: University of California, Los Angeles, Graduate School of Education and Information Studies, National Center on Evaluation, Standards, and Student Testing, 2000.