# The Power of Testing

MATTHEW GANDAL AND LAURA MCGIFFERT

*Just as medical tests help diagnose and treat patients, rigorous and meaningful education assessments can help ensure the academic health of all students.*

Our society relies on testing. We expect tests to tell us whether our water is safe to drink, our cholesterol is too high, or the dishwasher we want to buy is the best value. These tests help us ensure our safety, take care of our health, and spend money wisely.

In medicine, no one seriously questions the connection between testing and appropriate medical treatment. A patient may present certain obvious symptoms, but before making a diagnosis, a doctor will routinely order a battery of tests to isolate the specific condition causing those symptoms. Once these tests have identified the problem, the doctor can offer a treatment plan.

We may complain about the quantity, cost, or inconvenience of these tests, but we do not question their basic value. We expect our doctors to stay up-to-date on the latest testing methods, and we demand that our health plans provide coverage for the tests that we need.

When it comes to education, testing holds the same power to bring about the result we all want—academically healthy students. Although some may raise legitimate concerns about the adequacy of some tests now in use, we should not discount the validity or utility of testing altogether.

Imagine that every child had an annual education checkup—a set of assessments created to measure agreed-on expectations. The results of this checkup would help teachers chart a course for individual student improvement. The federal No Child Left Behind Act offers exactly this sort of checkup to U.S. schools and teachers. By the 2005–2006 school year, schools in every state will assess students in reading and math annually in grades 3–8, and again before they graduate from high school. Science assessments in key grades will follow in 2007–2008. No Child Left Behind puts tremendous pressure on states to create new assessments within a tight timeline. At least 36 states will have to develop more than 200 new tests within the next few years to comply with the federal law. If the end result is quantity without quality, little value will be added. If these assessments are of high quality, however, they have the potential to add significant value to school improvement efforts.

## HIGH EXPECTATIONS FOR ALL

Doctors treat each patient differently on the basis of his or her individual needs, yet they base their judgments on conventions widely held across the profession. A thermometer is the same in Indianapolis or Miami, as are a blood pressure gauge and a scale. What do these measures have in common? Each reports results against a common standard. A fever is a fever, regardless of where you live.

Similarly, we should hold students to the same achievement standards regardless of their race, their socioeconomic status, or where they attend school. Unfortunately, this does not always happen. Teachers' expectations for their students differ; an A awarded in one school can mean something very different in another school. And students in disadvantaged communities are disproportionately held to lower standards.

Challenging all students to meet common standards should be non-negotiable. These standards must be more than just minimum requirements; they must be anchored in the challenging content and skills that students need to succeed. The highest-performing school systems around the world use this formula of common standards and assessments. Students in these countries routinely outperform U.S. students on international assessments, not because they have more talent, but because their schools expect more from them (TIMSS International Study Center, n.d.).

If these international comparisons are not convincing enough, we can find plenty of other evidence of the need for common, high standards. Too many students graduate from high school unprepared for the challenges that lie ahead. Increasing numbers of students at four-year colleges need remedial education in reading, writing, or mathematics. Employers tell a similar story: 34 percent of job applicants tested by major U.S. firms in 2001 lacked sufficient reading and math skills to do the jobs that they sought (American Management Association, 2001).

## IT MATTERS WHAT WE MEASURE

Useful medical tests must provide relevant and reliable information. A doctor would not order an X-ray to determine treatment for a sore throat, or a throat culture to treat a broken ankle. Doctors need tests that reveal information about a patient's particular condition.

Similarly, useful education assess-

ments must make clear what they measure, and they must measure what we value most. In other words, states must tightly align assessments and standards to provide valid and meaningful information to educators.

Many states have found it difficult to accomplish this goal. Such alignment requires states to rely less on off-the-shelf, norm-referenced tests. These tests are not well aligned with most states' standards, and they report results against a norm, or average, rather than show whether students have met standards.

But even states that have developed their own tests have had trouble measuring their standards well. Our analyses of more than a dozen state tests designed to support standards-based instruction found that many tests are unbalanced, over-sampling some standards and under-sampling others. The more advanced content and skills usually get short shrift. For example, Achieve's research has found that although most states' middle school math standards emphasize the foundations of algebra and geometry, more than 60 percent of the questions on their 8th grade tests dealt with computation, whole-number operations, and fractions.

In contrast, Massachusetts is an encouraging example of a state that has established a well-aligned system of standards and tests. The state's 10th grade exams in English and math are among the most robust assessments in the United States. They are based on clear and challenging standards, and they measure the depth and breadth of those standards well. Students must read and write thoughtfully to do well on the Massachusetts English language arts tests, and they must demonstrate their understanding of both basic and advanced mathematics to do well on the mathematics tests. These exams will count for graduation this year for the first time, and although this requirement has engendered some debate, most people in the state agree that the tests measure what matters most (Achieve, 2002a).

## Using Data to Inform Practice

The patient's examination is over, the relevant tests have been completed, and the results have come back. Now, with all the information in hand, the doctor can offer a diagnosis and prescribe treatment. He or she may pronounce that the patient is in perfect health, or may recommend antibiotics, physical therapy, or even surgery. The important question is what happens after the results come back.

Just as in medicine, assessments in education are a means to an end. Assessments provide information on where students and schools need to improve, and they may provide incentives for students and schools to make the necessary improvements. But tests alone cannot create improvement. Educators, parents, and students must do the work of raising student achievement.

To make this possible, schools must get test results in a timely and useful manner. Waiting six months to see how students scored is of little value in helping those students improve. In addition, states must provide the test data in a form that educators can understand and use, with a finer degree of specificity than just a number on a scale. For example, a particular score in phonemic awareness conveys more to a teacher than an overall score in reading and certainly more than a score in English language arts. Specific results that identify students' particular strengths and weaknesses enable teachers to target instruction to meet the needs of each student.

New York City has dramatically altered how it reports results on state assessments in order to make them useful in classrooms. Partnering with the Grow Network, the city provides every parent, teacher, and principal with clear reports and instructional tools linking the data to state standards. Innovative technology can disaggregate the data to the individual level, allowing teachers to identify which students need help with which concepts, instead of requiring the entire class to review all topics. This approach has met with remarkable approval from educators and parents, who now find the data from the state tests much more useful (Grow Network, n.d.).

Meaningful data turn a diagnosis into action, thereby enabling educators to respond to individual student needs. They also make assessments a helpful tool for educators rather than simply an accountability hammer.

## Beyond Large-Scale Tests

Doctors routinely pair their own clinical observations with an objective test—like a blood cell count—to identify an illness. This practice allows them to get the most coherent and complete information and make an accurate diagnosis.

Educators, too, get the best information about their students when they compile data from a number of sources, including classroom assignments, quizzes, diagnostic tests, and large-scale assessments. Together, these tools paint a fuller picture of student performance than a single assessment can.

Large-scale state tests play a crucial role in monitoring and encouraging school improvement, but they are not enough. To tap the power of testing, schools and teachers need access to diagnostic assessments that give them immediate feedback on student performance throughout the school year. As states add new large-scale tests to meet the requirements of No Child Left Behind, school districts have the chance to drop duplicative tests and invest instead in diagnostic tools. Spring Branch Independent School District in Texas has done this to great effect by developing lessons, quizzes, and tests directly aligned to the standards that teachers can use in their classrooms at any time (Achieve, 2002b).

## The Real High Stakes

With the results of medical tests in hand, doctors have an ethical duty to give their patients the best possible care, regardless of the complexity of the disease. In fact, they can be held responsible if they do not provide appropriate treatment.

We, too, must provide the best education for all students – not just those who are easy to educate. Indeed, we are delinquent if we pass students through the grades and award them diplomas even if they are unprepared for the opportunities and challenges that await them. The real high stakes for these youngsters will come when they arrive at college or the workplace and lack the skills to succeed.

Like doctors who do not act responsibly, schools whose students consistently fail to meet expectations should face consequences. Rather than just a heavy stick, these consequences should include a combination of assistance and sanctions. In Kentucky and North Carolina, for example, the state assigns teams of distinguished educators to help low-performing schools develop and implement improvement plans that often focus on boosting the ability of teachers to teach to the state standards. Both states have seen a dramatic decline in the number of low-performing schools (Mandel, 2000). But ultimately, if assistance does not lead to improved performance, states must take stronger actions – including reconstitution or state takeover.

## The Challenge Ahead

Standards, testing, and accountability have become the policy framework within which schools in every state must operate. For schools and students to reap the benefits of standards-based reform, we need clear and rigorous standards, assessments aligned to those standards, results reported in meaningful ways, and appropriate incentives and consequences. States and districts also must work in tandem to align curriculum, diagnostic assessments, and high-quality professional development for teachers.

A few critics will always condemn the use of testing in schools. However, with students' futures at stake, we must not abandon the very tools that have the power to transform teaching and learning. We must make our education assessments stronger and take advantage of the information they provide to ensure that all of our graduates are academically healthy.

## References


Achieve, Inc. (2002a). *Three paths, one destination: Standards-based reform in Maryland, Massachusetts, and Texas*. Washington, DC: Author.

Achieve, Inc. (2002b). *Aiming higher: Meeting the challenges of education reform in Texas*. Washington, DC: Author.

American Management Association. (2001). *AMA survey on workplace testing: Basic skills, job skills, psychological measurement*. New York: Author.

Grow Network. (n.d.). *About the Grow Network* [Online]. Available:www.grownetwork.org

Mandel, D. R. (2000). *Transforming underperforming schools: A strategy for Tennessee 2000*. Berkeley, CA: MPR Associates.

TIMSS International Study Center. (n.d.). *Third international mathematics and science study—1995* [Online]. Available: http://timss.bc.edu/timss1995.html



*Matthew Gandal is Executive Vice President of Achieve, Inc., an organization created by governors and corporate leaders to help states improve their schools and raise academic standards; www.achieve.org.*

*Laura McGiffert is Director of Achieve's Mathematics Achievement Partnership.*