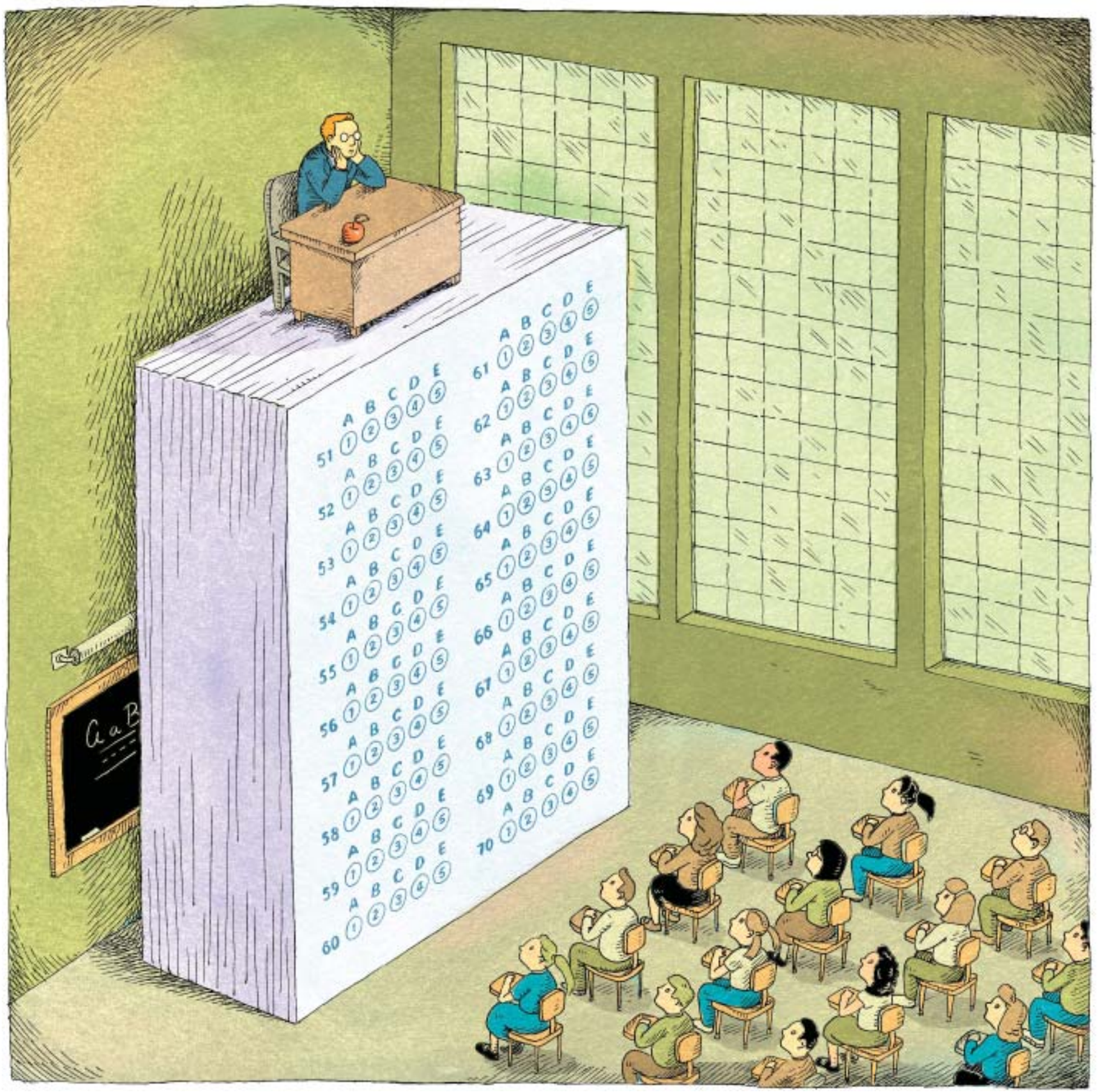

By Laban Carrick Hill

COMPROMISE IN THE CLASSROOM

Do new standardized-testing policies limit a teacher's ability to educate students properly?

Illustrations by Wesley Bedrosian

We are a country obsessed with measuring. If we can't quantify it, then a value cannot be placed on it. In a market economy, something without a specific value cannot be put up for sale and cannot be bought. In short, it cannot be a part of the economy. The way President Bush's 2001 "No Child Left Behind" (NCLB) Act is designed seemed to be based on these simple truths, which also suggests why it has received unprecedented support. The bill passed the House of Representatives by a 381 to 41 margin and the Senate 87 to 10. The No Child Left Behind Act became law Jan. 1, 2002. At the core of the NCLB legislation is standardized testing in reading, math and science to quantify the success or failure of America's schools. This kind of testing is meant not only to measure student progress, but is also an incentive to change the culture of the nation's schools.

The belief that our schools have failed in teaching our kids what they need to know has been increasing in recent years. According to the U.S. Department of Education, 88 percent of the United States public supports raising standards and requiring graduation exams, which suggests a real belief that schools are not performing. Beginning with the *Nation at Risk* report nearly 20 years ago, there has been a vigorous national debate over how to improve our nation's schools and our children's achievement.

The NCLB Act responds to these concerns by attempting to give schools and districts greater flexibility and control, to require only scientifically proven teaching methods, and to hold schools accountable for results. NCLB is the largest education law passed in more than 35 years, when Congress passed President Lyndon Johnson's Elementary and Secondary Education Act. The NCLB law increased education spending to more than $22.1 billion—a 27 percent increase over 2001 and a 49 percent increase over 2000 levels. "These reforms express my deep belief in our public schools and their mission to build the mind and character of every child, from every background, in every part of America," President Bush said during his first week in office in January 2001.

At the center of the law is a big stick called "Adequate Yearly Progress" (AYP), an ambitious program set up to hold school districts and schools accountable for their students' performance in the core content subjects of reading, math and science. While AYP does not tell states what standards their students must meet, it does insist that states create clear guideposts that students, teachers, parents and administrators can measure for academic achievement. The NCLB Act requires states to create annual assessments that measure what children know and can do in reading and math in grades three through eight. One of the innovations of this law is the insistence that states provide data on student performance by poverty levels, race, ethnicities, disabilities, and limited English proficiencies to ensure that no child—regardless of his or her background—is left behind.

While on the surface this new act seems not only admirable, but also badly needed, many education professionals already have questioned the innovation of its provisions. Charles B. Myers, emeritus professor of social studies education at Peabody College, finds himself a bit cynical when he thinks about the NCLB Act. "We've been developing standards of accountability for social studies teachers and colleges and universities teaching social studies teachers for five years," explains Myers. "This is nothing new."

Myers also questions the effectiveness of standards when, in many school districts, having a certified teacher in the classroom is a luxury. In places like California and New York City, where dramatic teacher shortages have existed for years, sometimes the most a district can hope for is a warm body in every classroom. For the current 2002–03 school year, New York City has more than 8,000 teachers who are not certified in the classrooms. It is clearly unrealistic for the New York City School Board to insist on full certification if the consequences are classrooms with no teachers. Myers recalls talking to a school principal from a rural district. "He had two openings in his school and advertised widely for teachers. None of the applicants who applied for the job was qualified to teach, but he had to fill those slots. He hired two teachers whom he never would have hired if he were not desperate."

No amount of certification requirements will help children learn if no teachers are available to do the job. Myers suspects the tough new teacher standards are simply a new way to avoid funding education at the levels necessary to attract good teachers. "Paying for testing is a lot cheaper than actually spending the money needed to improve schools," says Myers. If a school cannot find a certified teacher because there are none, the best standards in the world will do no good. "There have to be teachers to hold accountable, and that is just not the case everywhere."

The program at which Myers levels his most withering criticism is NCLB's "Adequate Yearly Progress" program, which requires schools to give students achievement tests every year from third through eighth grades. "Testing isn't going to make poor students better or make teachers more effective. Yearly testing is just a bigger hammer to hit teachers with," argues Myers.

Alecia Ford, MEd'97 (Peabody), who has taught sixth grade at Jere Baxter Middle School, an inner-city school, and now teaches at Meigs Magnet School in Nashville, confirms Myers's conclusions about student achievement. "If you were to grade my ability as a teacher by standardized tests, you would have given me a failing grade at Jere Baxter Middle School and an 'A' at Meigs Magnet School, even though I as a teacher did not change. At Baxter the biggest problem was getting the kids to school and getting them to focus in the class. They had so many social and economic issues outside school that dominated their lives that whatever was

While "Adequate Yearly Progress" does not tell states what standards their students must meet, it does insist that states create clear guideposts that students, teachers, parents and administrators can measure *for academic achievement.*

One of the other truths that standardized testing does not address is that many students who fall behind do not catch up in high school. Instead, they drop out. *In fact, the pressure of these tests can help to push these students out even faster.*

According to Myers, one of the other truths that standardized testing does not address is that many students who fall behind do not catch up in high school. Instead, they drop out. In fact, the pressure of these tests can help to push these students out even faster. Myers has found that happening in the classroom could not compete. Most issues I dealt with in the classroom were social, not academic. At Meigs, on the other hand, there have been no problems like that. The parents are involved and focused on preparing their kids at home for school. The kids are ready and willing to learn without real-life distractions. The peer pressure at Meigs is focused on learning."

Myers uses the example of his wife, who taught first grade, to illustrate another problem standardized testing cannot adequately measure. "When my wife taught first grade in an inner-city elementary school in Nashville, she had 25 or 26 children in her class. By the end of the year, at least 18 of those children who had begun the year in her class had moved away and had been replaced by new students. In addition, more than 20 of those children lived for some time during the year with an adult in the home who was not a parent." Myers makes this point because he says the two factors that influence student achievement most are the teacher in the classroom and the child's life at home.

it is not unusual for 75 percent of a class entering an inner-city high school to test below level, and for this same class to have a graduation rate of 40 to 80 percent. On the surface it appears this class has made great strides to catch up. In reality, the majority of failing students have dropped out because they have no chance of passing the standardized graduation exam. In this case the standardized test functions to discourage students from staying in school. Again, Myers sees a lack of funding beyond creating and administering the standardized tests. His fear is that preparing students to pass these tests will become the primary education goal, not the retention of those students falling behind.

When it comes to what happens in the classroom, Carolyn Evertson, professor of education in the Department of Teaching and Learning at Peabody College, has spent her career working with teachers on classroom management. Her work is concerned with creating conditions in which children can learn. Over the years she has seen new programs come and go with varying degrees of success. "I've seen this happen so many times that teachers can get overwhelmed by the sheer amount of work it takes to make learning a priority," she explains. "These new programs may last a couple of years and then something new comes along. Some teachers manage to work new trends into the classroom the best they can. Other teachers just have to start from scratch and restructure, and still others may ignore it."

"If we add a new element," Evertson worries, "the assumption of policymakers is that it's just additive, like beads on a string. You're just adding another bead. In reality it is much more complicated. A new mandate can add a geometrical complexity to teaching, especially if teachers are trying to create a classroom in which students have ownership in what they do, or if some children have different educational plans from other students. I'm thinking particularly of special-needs kids as well as gifted kids. We've known for a long time that for all students to succeed, we have to differentiate so that children have access to learning in the ways they learn best."

Anne Marie Elkins, MEd'95 (Peabody), a sixth-grade teacher at Isaac Young Middle School in Hendersonville, Tenn., knows exactly what she has to sacrifice because of increased standardized testing. "I'm a literature teacher," she says. "If a child is going to continue to love reading, the sixth grade is the time to help solidify that. It gets a bit old to keep hearing that what I need to be doing is teaching skills and essentially bore them to death." Elkins is a proponent of the Accelerated Reading Program in her school, which provides real incentives for reading. She worries, however, that this successful program is being crowded out of the curriculum because it is not measured on a test. She argues that a love for reading is a skill that will enrich students for their entire lives, rather than for a short-term goal like a test. For her this is the kind of sacrifice education is making for higher test scores.

Elkins explains that at the end of the year, standardized test scores provide the basis of her evaluation as a teacher. Whether her students gained an enthusiasm for reading that will carry them through life, whether children read Greek mythology, is not important to the way she is measured.

Evertson tries to put the potential troubles into perspective. "Suppose you're a teacher who has worked out a way for those who need more help in reading to get that help, and the kids who need more help in math to get that help. And suppose you've also put together groups of students working together to help each other, along with a whole-class situation in which there is some direct teaching. All of this is orchestrated throughout the day through a system of procedures that have become routine. Then there's a new mandate in which the focus is for all students to have high scores on achievement tests. This new emphasis doesn't just add a bead to the string. It adds a higher level of complexity that has to be sorted through in all those different ways in which each class functions."

Evertson's concern arises out of years of working with and observing teachers in the classroom. She cannot overemphasize how complex it is first to try to meet the needs of children who are much more diverse than in the past and then to do justice to a mandate no matter how wise or how needed it is. "It's just an organizational complexity that is very hard for teachers," explains Evertson.

She uses an example of a friend who had to integrate a new reading strategy mandate into her classroom in the middle of the school year. "By the time the new policy came down, she already had each child on a reading level matching both where they were and where they needed to go. She had managed to individualize her whole class for reading, which is extraordinary. The students were doing well and enjoyed reading, and they could move to more difficult books. Each child had his or her own plan. Then the directive came down that each teacher must create three reading groups in the classroom. This teacher just decided to divide the kids up in three groups while still utilizing their own plans so that when someone came to see if the directive was implemented, what he saw was three groups."

Though the new mandate may have been smart policy, to ask this teacher to abandon a strategy that was obviously working in order to conform to the new mandate did not simply mean more work for the teacher, but also potential harm to her students' learning. It categorically dismissed her classroom decisions and attempted to impose a new order that contradicted what was already working.

Business management consultants John Maleyeff, an associate professor at Rensselaer Polytechnic Institute in the Lally School of Management and Technology, and Frank C. Kaminsky, professor emeritus in the College of Engineering at the University of Massachusetts–Amherst, echoed Evertson's concerns recently in an editorial in the *Hartfort Courant*. Like Evertson and most every other educator, they support the use of standardized testing in schools.

Maleyeff and Kaminsky, however, "have serious concerns about the way tests are being implemented and the way the results are being interpreted." They stress three important rules of effective quality management that must be adhered to if the testing is to be effective. "First, always use statistical methods to distinguish between random variation in performance outcomes and real changes that may occur. Second, always use performance outcomes to understand and improve the system. Third, never use performance outcomes simply to reward or punish individual employees." They stress that management by fear encourages teachers to find ways to beat the system, and they suggest that programs like "Adequate Yearly Progress" are setting themselves up for disaster because they emphasize rewards and penalties without giving any real power to the teacher who is ultimately rewarded or punished.

"Consider the case in which the annual bonus of a teacher is based on the performance of students on standardized annual tests," they write. "On the surface, the policy appears to be a good idea. But this policy violates key rules of quality management. That teacher has no control over the quality and makeup of the incoming class, no control over whether the current year's test is similar to the previous year's test, and no control over random statistical variations." The consequence of this kind of quality management is a system that rewards "teachers and administrators who divert their attention from other subjects to spend an inordinate amount of time teaching to the test." Maleyeff and Kaminsky conclude that relying on this system is no better than rewarding teachers and schools on the basis of a coin toss.

Peabody graduate and schoolteacher Alecia Ford likes to make the distinction between constructive uses of statistics and punitive abuses. She cites a recent article in the *Tennessean* newspaper that listed all the scores of local schools and broke down the schools into economic class and ethnicity of their students. "The problem I had with that article was that it made certain schools and teachers appear bad. Statistics are helpful to informed educators because they can use the information to respond to those who are not doing well. But to simply label a school

or teacher bad is not helpful at all."

Recent graduate Elizabeth Amy Bantly, BS'02 (Peabody), who teaches sixth grade at H.G. Hill Middle School in Nashville, feels this public accounting is good for schools and communities. "We have to demonstrate quality to the outside world," she explains. "The public is not there with us every day, and they don't know what we're teaching and how the students are learning, so [standardized testing provides] an outside indicator of these things." In her first year of teaching, she has not felt a lot of pressure to completely change the curriculum. She remains optimistic, saying, "I think they want us to do what's best for the kids."

Still, Evertson worries because a standardized test doesn't measure everything a school teaches. "People who believe achievement tests are valueless really miss the point," she explains. "It truly is important, for example, that third graders have certain basic skills such as number sense, word recognition and vocabulary, and testing can serve a purpose in determining this. But, conversely, not performing well on the test doesn't necessarily mean that students are not going to do well in school. Other things come into play. Kids who don't do well on tests, but learn through different modalities, may not have the opportunity to use those modalities on the test. Let's say I understand a concept if I see it represented in pictures. If I am tested with words, I may not do very well even though I know the basic concept. I won't be able to show my understanding in a text-based mode.

"Tests are useful, but limited, tools for assessing what students know. They typically cover a pretty narrow range of skills. And if they are given the proper weight in the larger picture of student assessment, they provide important information (when combined with other things) to help teachers understand what needs to be taught or retaught. But not all students are able to show what they know through formal testing. It's when the tests become high stakes that it becomes a problem. Even the [college-level] GRE and SAT exams make allowances for taking the tests orally; I haven't heard about much of that kind of accommodation [in the AYP program]. The assumption is that kids should be able to take these tests."

The American Federation of Teachers … insists that teachers participate in the creation, implementation and evaluation of standardized tests, *that sufficient funding is provided for struggling students, and that the tests are used only for the purposes for which they were designed.*

Bantly concedes that Evertson's concerns are valid. "For some students testing is a good representation of their abilities, but with others not so. I have some students who have learning and attention difficulties, and I know they are a lot smarter than the tests. They might test on a third-grade level, but I know they can read on a sixth-grade level."

Charles Kinzer, former professor of teaching and learning at Peabody College and now at Columbia University, finds the insistence on one kind of assessment troubling. "There are lots of ways to assess," he says. "Teachers assess when they grade homework. Teachers assess when they see people on the playground sitting by themselves and not interacting with other kids. If they're not socialized into the classroom, they're not learning as well as they could. So there are lots of different kinds of assessment going on, including standardized. What people forget is that there are lots of goals and needs for assessment, such as the needs for a school board that is going to be ordering textbooks. The kinds of scores and information they need is different in some ways than what the classroom teacher needs to know for the next day's lesson. So there are different needs of assessment across the educational endeavor. The concern I have is that the act doesn't deal as much with teaching as it does with assessing. If we start putting a limited pie of resources into additional assessment, then some of the good things that could be happening in terms of instruction will go away." Though Kinzer admits it is too early to tell if this will actually happen, he feels strongly that these issues must be raised early.

"So much of what you hear in the popular press—and that's where perceptions get formed on the part of parents—is that we're going to do state tests and we're going to do national tests and we're going to make sure that no child is left behind because we'll be able to figure out who is not learning," Kinzer says. "But there is very little in the act that addresses what to do once we figure out who's *not* learning. We already know that in fourth grade, we have this tail-off. Some children do really well in the second grade and third grade, but when they hit the content areas, they don't learn as well as they could. That transition from 'learning to read' to 'reading to learn' gets stalled. They have learned to read stories OK, but when they get hit with textbooks, they tail off. The NCLB Act is good at identifying effective early reading strategies, but when kids hit the fourth grade and fifth grade, the act provides no direction. It can't be more of the same stuff. That obviously didn't work. All the act does at this point is test to see who's reading in fourth and fifth grade. What if they're not? Then what?"

Myers suggests that teachers and schools already know who is not performing and that standardized tests, therefore, provide redundant information. He sees the increased testing as an unnecessary and inadequate effort to improve student learning. "These tests are really good at assigning blame, but not particularly successful at solving the problem." He would like to see more of these funds diverted into solutions.

One bright spot in the NCLB Act is the work that Lynn Fuchs, professor of special education at Peabody College, has done for the "Reading First" program. Reading First is a $900 million state grant program that promotes the use of scientifically based research to provide high-quality reading instruction for grades K–3. "I served on the reading assessment committee for developing guidelines to examine the technical features of reading tests," explains Fuchs of her work for the U.S. Department of Education. "That process is one that states need to use as they incorporate assessments into their Reading First applications [for grant money]. I was also on the assessment and instruction committees for Early Reading First, a program for preschoolers. These committees are shaping what states need to do as they develop their applications for Reading First money."

Fuchs, who is also a research-program director in Vanderbilt's John F. Kennedy Center for Research on Human Development, is excited about the work she has done because this program will give school districts the kind of information they need to make good choices. "School districts can look at the available methods out there that they can invest their resources in, and have the information about which of those procedures produce which kinds of effects and which of those procedures there's just no information for," she explains. "That can help school districts spend their money wisely, both in purchasing materials and the money they spend in professional development. Generally, I think the move to try to put into place evaluative criteria of educational practices that rely on scientific evidence is a good thing, not a bad thing. If I had a child of school age, I would rather have my child in a classroom for which the school district has invested its dollars in methods for which there is scientific evidence showing it produces a good outcome."

Ironically, Reading First and Early Reading First hold the most promise for students, teachers, schools and parents even though its emphasis is on preschoolers through third graders. The majority of the NCLB Act focuses primarily on third through eighth grades and does not provide the kind of support that the Reading First and Early Reading First programs offer.

Recently, the American Federation of Teachers (AFT) passed a resolution on standards-based assessment and accountability that attempts to provide guidance for the AYP program. In the resolution the AFT stated it supports "appropriate, high-quality testing" but "continues to oppose the abuse, misuse and overuse of standardized testing." The organization also insists that teachers participate in the creation, implementation and evaluation of standardized tests, that sufficient funding is provided for struggling students, and that the tests are used only for the purposes for which they were designed. Finally, the resolution asks that other kinds of student evaluation be used on an equal basis. In short, the AFT supports standardized testing as long as it is not the only evaluation tool.

Marc Bernstein, president of Kaplan K12 Services and a former math teacher and school superintendent, argued in a letter to the *New York Times* editorial page that "teachers are handicapped when too much emphasis is placed on standardized testing." He wrote that "a narrow curriculum clips the wings of creative, enthusiastic teachers. There is a way to maintain standardized state testing … and to keep teachers reaching toward their 'Mr. Chips' potential, too. Use tests in combination with other evaluations. By having teachers assign a practical demonstration of student knowledge, like a term paper or oral presentation, we get a more complete understanding of a student's skills, underscore the value of teaching subjects not measured on state tests, and keep teacher motivation high."

"I think one of the issues state education leaders are struggling with," cautions Fuchs, "is what they use for their statewide assessment, what they can use for tracking annually, and how those assessments align and don't align. To some extent, the people in Washington are trying to figure that out, too."

Finally, Fuchs tries to put the new act into perspective as its first-year anniversary passes. "With any major initiative there are likely to be good things that come from it and unanticipated outcomes that are not what you hoped for," she says. "Who knows on balance how this will go, but I think it's worth trying to infuse the educational system with standards that help school districts spend their money wisely. If you look at Title One, for example, there has been an enormous amount of money spent on things with no evidence that they work. As a taxpayer, I think people in this country want to feel that their money is being spent wisely, even though there are not clear-cut answers. We would like to know in the meantime that the best information we have is being used." V

Laban Hill is currently working on a reader's guide to Jonathan Franzen's Pulitzer Prize-winning novel, The Corrections, *and is writing a cultural history of the Sixties. His cultural history of the Harlem Renaissance,* Harlem Stomp!, *was recently published by Little, Brown. He lives in Vermont with his wife and two daughters.*