









Throughout the 1980s, the call for higher achievement grew beyond the federal government and academia, spurred by the changing economy. In the early 1980s, the country was struggling against a recession and unemployment that went as high as 9.7 percent (Bureau of Labor Statistics, 2004). Powerhouse companies in Japan and Europe were competing successfully with American companies and, it seemed, jeopardizing our premier role in the world economy. Traditional well-paying jobs were disappearing and many people came to believe that a high-wage economy required a focus on “working smart”—that is, shifting away from jobs in which a strong back and willingness to work were all that was needed to make a good start in America. Not surprisingly, the weight of the business community got behind major education reform.

Along with the push for global competitiveness, increasing attention was being paid to educational equity. The huge achievement gap between black and white was becoming increasingly obvious. As one example, in October 1977, when Florida sophomores faced a functional literacy test that was a new requirement for a diploma, 78 percent of black students—but only 25 percent of white students—failed (Debra P. et al., 1979). And Florida was not alone. According to the National Assessment for Educational Progress (NAEP), throughout the 1980s, black 12th-grade students' scores in reading and math were about equal to those of white 8th-graders—and Hispanic students were not faring much better (NCES, 2000).

The growing perception among employers and higher education professors that the high school diploma had lost its luster; the nervousness about what all this would mean to our ability to compete in the increasingly global economy; the dramatic achievement gaps—all of these contributed to the growing belief among governors, policymakers, business leaders, and Americans generally that something had to be done to dramatically lift the quality of American education. By the end of the 1980s, many researchers and policymakers were beginning to converge on a solution.

### **The Promise of Standards-Based Reform**

In the 1980s and early 1990s, while dissatisfaction was continuing to build, some policymakers and researchers looked overseas at the education systems that had performed well on a variety of international assessments (Resnick and Resnick, 1985). Virtually all had education systems that were anchored by a national or nationally coordinated curriculum, which outlined in some detail the content and skills that students were expected to learn. Typically, students across these countries studied a common curriculum through at least 4th grade (Germany) and often through 8th or 9th grade (France), with students then streaming into separate educational tracks.

The existence of the national curriculum allowed for the creation of an entire education system geared to helping teachers teach the curriculum well. Teacher preparation and ongoing professional development were powerful because they were tightly focused on helping teachers understand the material they needed to teach and how to teach it. In most of

these countries, *examinations* given toward the end of secondary schooling were based directly on the national curriculum or publicly distributed syllabi. Publishing companies planned their textbooks and supporting materials around the specific syllabi and curricula. Finally, the curriculum and syllabi themselves were typically easily available to the public; it was even for sale at regular bookstores. As a result, students, parents, and teachers all knew what kids should be learning; the possibility that expectations for poor and affluent students, especially in the lower grades, would be quite different was greatly diminished (see "[Lack of Equity, Quality Push Standards Forward in '90s](#)").

## **II. An American Educational System Based on Standards**

Americans liked the coherence, alignment, and achievement results of these systems, but their centralization grated against the American tradition of local control of schools. The search was on to find a uniquely American way to capture the benefits of an aligned education system, without losing local control. A great national discussion ensued. Among the strong public voices advocating an education system driven by clear, high, transparent academic standards was AFT's president, Albert Shanker, who wrote on the issue many times in his weekly *New York Times* column (see box above). An influential paper by Marshall Smith and Jennifer O'Day (which began circulating long before it was published in 1991) described a potential American version of such a steering system.

In 1989, the discussion moved to the top of the American agenda, when the National Governors Association (NGA) hosted the President's Education Summit with Governors. The groundbreaking meeting endorsed the idea of national educational goals and a process for pursuing them that didn't undermine local control. From there the discussion moved to the newly established, bipartisan National Education Goals Panel (whose Resource Group on Student Achievement was chaired by Lauren Resnick, this article's lead author), and then to the congressionally authorized National Council on Education Standards and Testing, which included elected leaders from both parties and private individuals from the worlds of education (including Lauren Resnick), business, and other fields. Following much debate, discussion, and compromise, a rough consensus emerged, as captured in the documents produced by these various groups, on the main elements of what has come to be known as standards-based education. The basic tenets, which were further developed and honed in states and in federal legislation, included the following:

1. Use a public process—involving educators, parents, community members, and potential employers—to establish common and transparent expectations, known formally as standards, for what students should know and be able to do upon graduation and at certain key earlier grade levels.
2. Develop assessments geared to standards that students could prepare for and that could provide clear targets for teachers' instructional work with students.

3. To preserve local control, encourage districts and schools to enact instructional programs explicitly geared to the standards and to organize continuing professional development around those programs. Pre-service teacher training, too, was to be organized around the standards.

4. Create accountability systems that are based on whether students are meeting the publicly set and assessed standards.

The idea was that a standards-based system could combine the positive aspects of centralized curricula with the individuality and energy of the American local control system. The standards and assessments would be set by public entities such as states, but the details of curriculum, teaching, and professional development would be left to districts and schools. The accountability systems, rather than detailed regulations, would structure the priorities of schools and districts and press them to make the changes necessary to deliver effective teaching to all of their students.

It was an imaginative effort to harness the power of alignment without diminishing local control. It's also now clear that the task left to schools and districts—to create their own curriculum and instructional programs *and* figure out how to reinvent themselves to effectively deliver those programs and do it quickly—was enormous. The capacity of the schools to dramatically improve education was quickly outpaced by the much faster moving development of assessments and accountability systems. And this created the difficulties mentioned earlier: the inadequate support for teachers to meet ambitious new educational goals, the excessive focus on test preparation—in fact, in many places, the virtual hijacking of standards and education by narrow tests.

How did we get here—and how can we get back to the original intent of the standards-based system? To answer these questions, we'll look first at the development of standards; second, at the difficulties that have been confronted in bringing standards-based assessments to life; third, at the ways in which curriculum and professional development have (or haven't) been built around these standards; and fourth, where the rubber really hits the road, the accountability rules brought to us by the No Child Left Behind Act, signed into law in 2002.

### **Standards**

We begin with the academic standards. Who would write them? How detailed would they be? The years following the National Governors Associations' (NGA) 1989 Summit were a time of ferment as states, associations of states, the federal government, professional societies, non-profits, school districts, and individual schools all set about writing standards. For a time, it looked as though the lead role might go to national professional associations, such as the National Council of Teachers of Mathematics (NCTM), which in 1989 had written the first set of home-grown national standards. But there were also large states that developed their own standards or "curriculum frameworks"; one of the first in this early generation of standards documents was California, which launched a set of curriculum frameworks starting in 1987.

And, there were efforts, such as that of the New Standards Project (which Lauren Resnick co-directed with Marc Tucker) to bring together consortia of states to prepare standards and related assessments (Viadero, 1994).

The first clear inkling that states would end up as the main makers and adopters of standards came in 1994. In that year, President Clinton signed the newly revised Elementary and Secondary Education Act (ESEA), renamed the Improving America's Schools Act, which required states to set statewide academic standards for its Title I students *that were the same as the standards that existed for other students*. This, of course, required any state that had not yet adopted standards to do so. The trend toward state standard-setting was locked in when No Child Left Behind (NCLB) became law in 2002. NCLB was yet another revision of ESEA, this one backed by President George W. Bush, with bipartisan congressional support.

Though the debate over who should set standards had subsided, there remained a question of just what standards should look like. How general? How specific? Should there be separate standards for every grade or should standards be specified just for broad "grade spans," such as 1-4 or 5-8? States that chose to set periodic rather than grade-by-grade standards did so for what seemed to be a good reason—a widely held view among policymakers and others that we did not want people from outside the local school district to control every step of the curriculum. The same thinking led some states, whether their standards were grade-by-grade or periodic, to develop very general standards, instead of more detailed ones that approached the specificity of a curriculum (although there were also several states whose standards were so specific and lengthy that they were impossible to actually teach). But this view, that standards should be vague and/or periodic, ran into trouble, as we will see, as assessment and accountability developed.

While standards have gone through more than one round of revision in most states, they continue to vary widely in style and quality. Recent analyses of the overall quality of standards show a mixed picture and sometimes fail to agree on which states have good or bad standards (Stotsky and Finn, 2005; Klein et al., 2005; *Education Week*, 2004). (For examples, see box below.)

STRONG STANDARDS vs. WEAK STANDARDS	
SCIENCE	
Describe how groups of elements can be classified based on similar properties, including highly reactive metals, less reactive metals, highly reactive nonmetals, less reactive nonmetals, and some almost completely nonreactive gases. (grade 8)	Describe the historical and cultural conditions at the time of an invention or discovery, and analyze the social impacts of that invention. (Grades 5-8)
SOCIAL STUDIES	
Describe major rights, such as freedom of speech and freedom of religion, that people have under Indiana's Bill of Rights (Article I of	Students will trace patterns of change and continuity in the history of their community, state, and

the Constitution). (Grade 4)	nation and in the lives of people of various cultures from various periods. (Grade 4)
<p><i>Above are examples of "weak" and "strong" standards, as evaluated by AFT Educational Issues staff. The AFT's annual evaluation of state standards was published as Making Standards Matter, from 1995 to 2001. Since the late nineties, the AFT's reviews have been published annually by Education Week. For more information, including AFT's current state-by-state analysis of standards see <a href="http://www.aft.org/topics/sbr">www.aft.org/topics/sbr</a>.</i></p>	

Clearly, the low-quality of some states' standards is a major barrier to realizing the potential benefits of standards-based education. Good standards are the foundation for the other elements of standards-based reform: a rich curriculum that builds important knowledge and skills in a logical sequence, professional development that focuses on teaching the curriculum, and assessments that measure whether students are reaching the standards.

### **Assessments Aligned to Standards**

Standard-setting was the crucial first step in building a standards-based education system. Next, in terms of attention and importance, was a new function for testing and assessment. Standards-based assessments were meant not just to judge performance by students and teachers (an accountability function, which we will discuss later), but also to serve as guideposts for teaching and learning. The idea was to create assessments that students could prepare for and that teachers could legitimately prepare students to do well on.

The idea of assessments designed to be taught to and studied for was new to most Americans (though in New York State, the Regents exams were of this type). But it was an idea familiar in most other developed countries that had for decades been using public examinations both as a basis for granting secondary school certificates and for university entrance (Resnick and Resnick, 1990). In examination-based education systems, it is normal and appropriate that curriculum and teaching are related to exams and aimed at helping students do well on them. In most European and Asian countries, for example, secondary school students take subject-matter examinations that are directly linked to a publicly specified curriculum. In some countries the exams are graded centrally by teams of teachers; in others, teachers grade the exams in their own schools and a sample of papers are graded centrally in order to "calibrate" local scores (so that grades coming from different schools, or even different cities, are comparable and students everywhere benefit from common expectations).

A crucial feature of these examinations is that students are rarely surprised by them. Both teachers and students know what to expect, indeed teachers draw on past exams as instructional guides. Not everyone likes all of the questions and study tasks, but teachers and students view the system as fair. What is more, external exams of this sort have the effect of turning students and teachers into a "team," jointly working towards exam preparation. Similar teaching is seen in

the U.S., when teachers prepare students for such externally developed exams as the Advanced Placement, International Baccalaureate, or some statewide exams, rather than their own end-of-course tests (Resnick and Resnick, 1992).

Examinations of this sort can take multiple forms. They can be "on demand" assessments in which students respond to set questions, including multiple choice questions, short constructed responses, extended essays, "performance assessments," or extended pieces of student work produced over a longer period of time ("portfolio assessments"). An ordinary-looking test or an open-ended performance task becomes an examination when it is explicitly aligned to the curriculum or standards that students are meant to learn. Teaching toward well-constructed examinations is good professional practice.

Unfortunately, the tests that most states adopted to judge student progress toward state standards were not of this sort. Some states used the same tests (sometimes in adapted forms) that for years they had been purchasing from American testing companies, and these were not designed as exams. They were not systematically aligned to a specific curriculum or to standards that established what students should learn. Instead, they were designed to compare students with each other—spreading them out on a "bell curve." The most common way of describing how much students knew, based on these tests, was to declare their "percentile" scores. Typically, being "at grade level" simply meant that you were at the 50th percentile—half of the norming sample scored higher than you, half lower. To make the tests work this way, test developers collected large pools of items that were thought to sample the average curriculum in use in American schools, tried them out on large populations of students, and then performed sophisticated statistical analyses to pick out the items that best "discriminated" among students—that is, spread them out on a normal curve. This was a far cry from constructing a standards-referenced or curriculum-referenced exam, in which one started with what one expected students to learn and developed test questions (or performance tasks) explicitly to match the standards or curriculum.

Many teachers have objected to being pressured to teach to norm-referenced tests, and indeed teaching to these tests is a bad idea. They were not designed to be taught to. Because they were meant to be used by many different school systems using many different curricula, they were not aligned systematically to anyone's standards or teaching programs. In addition, because these tests depended on spreading students out on a curve, test items were retained or omitted in a test based on how they discriminated among students, not on how well they represented the standards to be taught. For all these reasons, it was impossible to tell from typical norm-referenced tests whether students were actually learning to expected standards.

Unfortunately, the problem of weak tests, not fully aligned to standards, is not limited to recycled versions of "off-the-shelf" tests. Even states that have constructed their own tests, based on their own standards, have largely relied on

traditional test items and low-cost methods of scoring. But in a standards-based education system, everything depends on how well assessments actually represent the full range of standards, in both topical content and cognitive demand—and thus on what kinds of teaching and learning behavior they evoke. Unfortunately, most state tests are not well aligned to state standards. In some extreme cases, alignment between state standards and tests is so weak that the standards from one state more closely match the tests used in another state (Porter, 2002). Most state tests do not do a good job of assessing the full range of standards and objectives that the states have laid out for their students. In fact, research has found that what “is included and excluded is systematic: the most challenging objectives are the ones that are under-sampled or omitted entirely... [and those] that call for high-level reasoning are often omitted in favor of much simpler cognitive processes” (Olson, 2003). As a result, although most state standards explicitly call for conceptual understanding and problem-solving, their tests often fail to assess these standards. When teachers match their teaching to what they expect to appear on state tests of this sort, students are likely to experience far more facts and routines than conceptual understanding and problem-solving in their curriculum.

One could argue that with these tests we could at least measure whether students were acquiring the basics and that, for some students, a concerted effort to assure that they acquire the basics represents an improvement. But, as we will see, with the addition of accountability—and without a curriculum that defines broader educational goals—narrow tests may not serve simply as a floor, but can become the de facto curriculum. In short, the tests can hijack the rest.

### **Curriculum and Professional Development Aligned to Standards**

A strict test-based accountability system invites this kind of test-matching behavior. In theory, it is the standards that teachers should be aiming for, but it is the far narrower tests that carry the consequences. Many principals distribute practice material designed to prepare students for the tests; and commercial test prep materials, billed as diagnostic and useable as a basis for differentiating instruction, can be bought easily from various publishers.

That kind of test-driven teaching was not the goal of the standards movement. According to the vision put forward by the Goals Panel and the National Council on Education Standards and Testing, school districts would develop rich instructional programs with strong content and good pedagogy that would be explicitly aligned with state standards (not tests). The system would be a *coherent whole*, its practical functioning boosted by ongoing professional development. Tests would be part of that whole, but they would be grounded in the standards. And the specifics of how students would be taught the standards would be left to local decision-makers.

This element of local decision-making was the major factor that made an American standards-based system different from the national curricula used by other countries. But is it

working? The goal, remember, was to produce the benefits of Europe's and Asia's nationalized curricula—a common, transparent curriculum for all kids; a basis for powerful, focused pre-service and professional development; quality textbooks and curriculum materials; and an assessment system that would enable teachers, parents, students, and the country to measure students' progress toward mastering the curriculum—without actually producing a national curriculum.

To realize these benefits, someone has to create a curriculum or standards specific enough to carry the load that is carried by other countries' national curricula. It could be the state, the district, or an independent group. But without a common curriculum to serve as the anchor, standards-based reform cannot produce the aligned system of professional development, textbook and curriculum materials that was promised.

The bad news is that, as Achieve noted in a 2002 report, most states have not provided teachers or others with clear curricular guidance. According to an earlier 2001 report by the American Federation of Teachers, *Making Standards Matter*, only nine states had in place even half of what was necessary to provide teachers adequate curriculum guidance.

At the district level, the news has not, until very recently, been inspiring either. Our Institute for Learning works with some of the urban school districts that are trying the hardest to raise their students' achievement. Our work often begins with a "stock-taking" that includes examinations of test data coupled with classroom visits and discussions with teachers aimed at understanding the ongoing teaching program. We ask, "What is your curriculum?" Until recently, in most of these districts, both teachers and administrators described their curriculum by naming a textbook. Further discussion revealed that rather than defining a coherent program or assuring a common curriculum for all, the textbook was treated as a resource from which teachers could pick and choose materials for lessons, often adapting the material for their students. Teachers often did not know what their neighbors—teaching the same grade, and the same course, and similar students—were doing with the adopted textbook, or even whether they were seriously using it. Consequently, students often experience a very fragmented program over the course of several years, a situation that is particularly negative when students (and even some teachers) change schools frequently. De facto, then, there often was no coherent curriculum, even within individual schools. Thus, the foundation of poorly aligned standards and tests is now overlaid with weak curriculums—leaving teachers and the educational system with no common anchor except the tests.

We see hopeful signs that things are beginning to change, however. A number of districts, especially urban districts with mobile student populations, are beginning to recognize that a common curriculum across schools is a necessity. To boost student learning, some districts are also realizing that they need to greatly strengthen professional development, giving teachers the knowledge and skills they need to successfully teach challenging student populations that in the old days

were expected to put in their seat time but not learn much. These districts are also realizing that effective professional development is based on a particular curriculum; it's not general and vague. In short, effective professional development requires the adoption of a curriculum; and the effective use of the curriculum requires ongoing, classroom-based professional development for teachers.

In response, many districts are going beyond merely adopting textbooks to implementing more fully "designed" curricula. They sometimes adopt programs designed outside the district (e.g., Open Court Reading or Everyday Math). Sometimes, they build district-wide instructional guidance systems that may use a textbook as a base, but add substantially more pedagogical guidance. These instructional guidance systems go well beyond the old "scope and sequence" charts that mainly suggested a flow of content. The new guidance systems can specify sequences of topics, suggest specific instructional practices both from a textbook and "supplementals" (or classroom libraries), the amount of time each topic should take, curriculum-embedded assessment tasks, student work samples, and sometimes, model lessons for use in professional development. Although these instructional curricula are sometimes tightly defined, all of those where we have seen achievement increases specify a mix of conceptual and skill emphasis. None call for teachers to read a "script" to students or to expect preprogrammed answers from them. All depend on providing intensive professional learning opportunities for teachers. These positive results of linking professional development to a specific teaching program are what we might expect given the growing body of research demonstrating that academic achievement increases when professional development focuses on the specific content teachers are expected to teach (Cobb et al., 1991; McCutchen et al., 2002). In one study, for example, David Cohen and Heather Hill (2001) found that most teachers who reported improved instructional practices had attended substantial training programs focused specifically on the curriculum materials that they used in their classroom. Those teachers' schools also posted higher scores on a state mathematics assessment. By contrast, other professional development programs showed no such effects.

What we see, then, is the beginning of an effort in a growing handful of districts to make standards-based reform realize its full vision, not just instruction narrowed to tests. And, where it is happening, student achievement seems to be responding. But the magnitude of the effort being exerted in these districts cannot be minimized. The work we've described is typically being undertaken in large urban districts with strong district leadership and community support, where the infrastructure and economies of scale exist to support the large-scale implementation of curriculum and related professional development. But even in these districts, assembling the resources and know-how has been a challenge. What will be necessary to help other less able districts to move in this direction? What about students in smaller districts that can't afford such an investment? States—probably with federal support of different kinds—are going to have to figure out how to bring curriculum guidance and professional development to

a much larger population.

### **Accountability**

And so we arrive at a discussion of accountability. We've seen that the standards that exist around the country are of mixed quality, with many quite weak and vague. Layered on top of these weak standards are tests that are typically not well-aligned; and in almost all cases, the tests measure students' progress on basic knowledge and skills, but rarely on the higher-level cognitive abilities that are included in the state's standards. In many places, there is no detailed curricular guidance that would allow teachers across a district to teach a common curriculum that went beyond what was tested; and without this curriculum, obviously there is not the related training that would support teachers in teaching it.

If you layer high-stakes accountability atop all of this, the formula is complete for allowing a narrow test, focused on the lower end of the curriculum, to hijack broader educational goals. And indeed, in many places, that was beginning to happen even before No Child Left Behind (NCLB). But with the adoption of NCLB, the threat became nationwide. Whatever pressure already existed to teach to the test increased, both because the consequences imposed by NCLB were more dramatic and—due to NCLB's formula for defining whether schools had made "adequate yearly progress" (AYP)—because they affected more schools.

Any accountability system layered on such a weak foundation would cause problems. But NCLB has unique features that cause additional, unique problems. Among them: Its formula for judging whether schools have made AYP does not take into account where a school started or how much progress it has made, which means that schools that have made great progress (but not enough to make AYP) will nonetheless be identified as "in need of improvement" (see "[The AYP Blues](#)"). The particular consequences that it prescribes and the order in which they are prescribed can mean that wrongly identified schools will be subjected to consequences that can impede their further progress and thus hurt their students.

Further, the requirement that everyone must be proficient by 2014, while meant to encourage states to set high expectations for all types of students, is in reality encouraging states to set lower standards for everyone: The lower the standard, the easier it is for schools to meet the targets and avoid sanctions. Leaving the standards up to states was, of course, among the political compromises that made NCLB possible. It is a "states' rights" and "local control" solution embedded in a national law. But it creates an incentive to lower, rather than raise, expectations. For example, Pennsylvania deliberately lowered its proficiency standards after too many schools failed to clear the AYP bar. Some commentators believe the current law is creating a "race to the bottom," undoing years of gradual rises in expectations and achievement (Ryan, 2004).

But even layered on a weak foundation, accountability, as it has played out, whether under NCLB or under certain state and local systems, has a silver lining that should not be

dismissed lightly. It has brought substantial attention to teaching core, basic skills to the lowest-performing students and to a variety of programs that are increasingly aimed at improving the lowest performing schools in a district (see "[Standards-Based Reform Brings New Attention to Key Elements Necessary for Improving Student Achievement](#)"). And, in the case of NCLB (and state and local systems that disaggregate test data according to minority and poverty status), it has brought a special spotlight, and needed instructional focus, to helping poor, minority, ELL, and special education students improve their performance on the narrow (but important) body of skills and knowledge defined by state tests.

Nonetheless, the goals of standards-based education were much broader and higher than this. If we want to realize the benefits of standards-based education for the full range of students, and if we want our lowest performing students to reach the high standards that were originally the hallmark of the standards movement, accountability that is so heavily tied to poor tests and that doesn't assure that teachers get the support they need to teach to the standards will not get us there.

### **III. Where Do We Go from Here?**

If we mean to realize the benefits of standards-based education for the full range of students, and if we want all of our students to reach the high achievement levels that were originally the goal of the standards effort, we will have to attend to more than tests and accountability systems. The nation's efforts to truly realize the goals of standards-based education will be frustrated by the incompleteness of the reforms that have been put in place so far. The instructional support system—curriculum, instructional programs, professional development, targeted interventions for struggling students—that now exists in most districts is not strong enough to produce achievement that goes beyond bringing the basics to a larger group of students. And, if we continue to neglect the core of the reforms while pressing forward on accountability, we will engender more and more hostility from the public and educators who want what standards-based reform promised: truly high standards for all. Each of the four key elements of the standards-based education system needs attention.

*Standards.* States need to strengthen the specificity and clarity of their standards so that these standards can adequately play the role of other countries' national curricula. Standards should be clear, specific, and comprehensive enough to serve as the basis for building both good examinations and strong instructional programs. Grade-by-grade standards seem to provide the best guidance. So do standards that specify the kinds of texts to be read, the particular scientific or mathematics concepts to be learned, and detailed and understandable criteria for good writing and other complex skill performances. The danger in a move to specificity—long lists of topical content or mechanical skills to be mastered—has made some educators wary of detailed standards. It is time to take a new look and to find a "Goldilocks" solution—a workable middle ground between too

much and too little detail in standards. In this difficult process, states may find it useful to borrow from one another or from existing published syllabi and standards, or to join consortia that are developing shared standards. An example is Achieve's Mathematics Achievement Partnership, a group of nine states working together to raise expectations and improve student performance in middle- and high-school mathematics. As part of the initiative, Achieve has published a framework for what American students need to know in mathematics in the middle grades (Achieve, 2001).

*Assessments.* There is much to do, as well, on the testing and assessment front. Assessments play a dual role in a standards-based system. They are instruments for monitoring and accountability and, at the same time, they inevitably model and guide instruction. The higher the stakes, the more educators will feel pressed to teach to the tests. Therefore, the higher the stakes, the more important it is that assessments guide educators, and students, toward the kind of learning we truly want. We have seen that most of today's state tests are not well aligned to standards and they are most likely to leave out the most intellectually challenging aspects of the standards. Yet, it is the tests rather than the standards that claim educational attention.

To recapture the intent of the standards-based system, most state assessments need to be redesigned so that they guide teaching in the direction really intended by the standards. This will probably require adding substantial numbers of tasks that require open-ended and constructed responses, as is the practice virtually everywhere else in the world. There is no mystery about how to do this in ways that meet technical standards of measurement. But there is no doubt that standards-referenced assessments that include substantial numbers of open-ended and constructed-response tasks will be costly. Substantial assistance from the federal government is likely to be needed by all but the largest states. Dollars granted to states or consortia for this essential work will help.

*Instructional programs and professional development.* With standards and assessments still needing substantial work, it is perhaps not surprising that instructional programs and professional development geared to standards are barely out of their infancy. Here, too, we will need "Goldilocks" solutions that provide guidance that is detailed enough so that teachers don't each have to invent their own program, while leaving enough room for adaptation to students. Recent research has made it clear that professional development works best when it is tied directly to the program that teachers are using with their students. Programs that require teachers to follow word-by-word scripts and extremely prescriptive time schedules are unlikely to engage the best minds and the most committed educators for long. But leaving teachers to guess at what are the best ways to teach does not work either. Again, the task is to find the right balance. But even the best instructional programs will fail with some students; the structure and resources must be available to provide these students with intensive interventions.

Where districts develop their own instructional and

professional development systems, “buy-in” may be greater. But only the largest school systems usually have the resources for full program development. States may need to provide more tools and direct assistance than they do now, as well as more financial resources. And there is a key role for the federal government in supporting the development and testing of the kinds of research-based instructional systems that we have referred to as “designed programs.”

*Accountability.* Forms of accountability that keep the education system focused on important academic achievement goals and on equity—providing a high-quality education to all of our students—are essential. As we have noted, many aspects of the current NCLB accountability requirements need to be adjusted. For example, ways need to be found to measure and reward achievement growth, thus taking into realistic account our schools’ different starting points. And we need to reconsider some of the accountability requirements for special education students and English language learners. Thoughtful individuals in the states, the federal government, and the research community are at work on these issues, and we remain confident that the accountability provisions of NCLB can be adjusted as we learn more about how it actually works.

However finely tuned the accountability rules, however, they cannot have their intended effect on the quality and accessibility of education unless the first three components of the standards-based system are brought up to par. The accountability aspect of the program is, if anything, running dangerously ahead of the system as a whole. Because the stakes are high, the incentives to match teaching to tests instead of standards are almost irresistible. And if we don’t sharpen standards and assess what we really mean by them, the nation is likely to wake up in a few years to find that it has created a “fool’s gold” system. We will have more and more of the least valuable coin of the realm; while the high levels of achievement we meant to create will increasingly elude us.

Much has been accomplished since the National Governors Association summit that put standards on the front burner. But, increasing student achievement beyond a relatively low standard will be nearly impossible unless we create the coherent whole that inspired the standards movement 15 years ago.

The original vision of standards-based education, we think, was the right one. And some notable progress has been made: Standards are now in place, although some will need substantial revision before they can adequately guide educators toward the intended high expectations for learning. Accountability has produced unprecedented attention to the very students it had been easy to ignore, or to set low expectations for, in the past. But standards and accountability are only the outer shell of the standards vision. The core of the reforms—aligned, high-quality assessments, instructional programs, and professional learning opportunities—have yet to be realized.

---

Lauren Resnick is director of the Learning Research and Development Center (LRDC) at the University of Pittsburgh, and founder and director of the Institute for Learning, which provides professional development to urban school districts. She is also editor of *Research Points*, a publication of the American Educational Research Association, and is co-founder and co-director of the New Standards Project, which has developed educational standards and assessments for states and school districts. Chris Zurawsky is LRDC's communications director and managing editor and an issue writer for *Research Points*. The authors wish to thank the National Science Foundation for partial support of the preparation of this paper. The opinions expressed in this paper are the authors' and do not necessarily reflect those of the Foundation. For more information on LRDC, visit [www.lrdc.pitt.edu](http://www.lrdc.pitt.edu).

Related Sidebars:

["An American Revolution: A Common Curriculum,"](#)  
by Albert Shanker  
["Lack of Equity, Quality Push Standards Forward in '90s"](#)

---

\* Since researchers have yet to agree on the proper way to calculate graduation rates (e.g., whether or not to include people who have earned a GED), readers have probably seen higher and lower graduation rates than these from NCES.

---

#### References

Achieve, Inc. (2002). *Staying on Course: Standards-Based Reform in America's Schools: Progress and Prospects*. Washington, D.C.

Achieve, Inc. (2001). *Foundations for Success: Mathematics for the Middle Grades*. Washington, D.C.: Mathematics Achievement Partnership.

American Federation of Teachers (2001). *Making Standards Matter*. Washington, D.C.: American Federation of Teachers.

Barnett, W.S., (2005). Personal communication.

Barnett, W.S., Huystedt, J.T., Robin, K.B., Schulman, K.L. (2004). *The State of Preschool: 2004 State Preschool Yearbook*. New Brunswick, N.J.: National Institute for Early Education Research.

Boyer, E. L. (1983). *High School: A Report on Secondary Education in America*. New York: Harper & Row.

Bureau of Labor Statistics (2004). *Employment Status of the Civilian Noninstitutional Population, 1940 to date*. Washington, D.C.: U.S. Department of Labor.

Center on Education Policy (2005). *From the Capital to the Classroom: Year 3 of the No Child Left Behind Act*. Washington, D.C.: Center for Education Policy Report.

CGCS (2005). *Beating the Odds: A City-By-City Analysis of Student Performance and Achievement Gaps on State Assessments. Results from the 2003-2004 School Year*. Washington, D.C.: Council of the Great City Schools.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., and Perlwitz, M. (1991). "Assessment of a problem-centered second-grade mathematics project." *Journal for Research in Mathematics Education*, 22, 13-29.

Cohen, D.K. and Hill, H.C. (2001). *Learning policy: When state education reform works*. New Haven, Conn.: Yale University Press.

Debra P., et al. v. Turlington, et al., 474 F. Supp. 244 (Fla. Dist. 1979).

Education Week (2004). Quality Counts 2004: Count Me In. January 8, 2004.

Jennings, J. and Hamilton, M. (April/May 2004). "What's good about public schools." *Our Children*.

Klein, D., Braams, B.J., Parker, T., Quirk, W., Schmid, W., Wilson, W.S., Finn, C.E., Torres, J., Braden, L., and Raimi, R.A. (2005). *The State of State Math Standards 2005*, Washington, D.C.: The Thomas B. Fordham Foundation.

Mandlawitz, M. Ed. (2005). *Education Budget Alert for Fiscal Year 2006*. Washington, D.C.: Committee for Education Funding.

McCutchen, D., Abbott, R. D., Green, L. B., Beretvas, N., Cox, S., Potter, N. S., Quiroga, T., Gray, A. L. (2002). "Beginning literacy: Links among teacher knowledge, teacher practice, and student learning." *Journal of Learning Disabilities* 35(1), 69-86.

Merriman, D. (2000). *What Accounts for the Growth of State Government Budgets in the 1990s?* Washington, D.C.: The Urban Institute.

National Center for Education Statistics (2003). "High School Graduates Compared With Population 17 Years of Age, by Sex of Graduates and Control of School: Selected Years, 1969-70 to 2002-03." *Digest of Education Statistics and Figures 2003*. Washington, D.C.: U.S. Department of Education.

National Center for Education Statistics (2000). *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance*. Washington D.C.: U.S. Department of Education.

National Commission on Excellence in Education (1983). *A Nation at Risk: The Imperative for Educational Reform*. Washington, D.C.: U.S. Government Printing Office.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, Va.: NCTM.

Olson, L. (2003). "Standards and Tests, Keeping Them Aligned," *Research Points*, 1 (1) 2003. Washington, D.C.: American Educational Research Association.

Olson, Lynn (2002). "Survey Shows State Testing Alters Instructional Practices." *Education Week*, Volume 21, Issue 32, pp. 14, April 24, 2002.

Powell, A., Farrar, E., and Cohen, D. (1985). *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*, Boston, Mass.: Houghton Mifflin.

Porter, A.C. (2002). "Measuring the Content of Instruction: Uses in Research and Practice," *Educational Researcher*, Vol. 31, No. 7, pp. 3-14.

Resnick, D.P. and Resnick, L.B. (1985). Standards, curriculum, and performance: A historical and comparative perspective. *Educational Researcher*, 14(4), 5-20.

Resnick, L.B. and Resnick, D.P. (1990). Tests as standards of achievement in schools. In G.R. Anrig (Ed.), *The uses of standardized tests in American education: Proceedings of the 1989 ETS Invitational Conference* (pp. 63-80). Princeton, N.J.: Educational Testing Service.

Resnick, L.B. and Resnick, D.P. (1992). *Assessing the thinking*

curriculum: New tools for educational reform. In B.R. Gifford and M.C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer.

Ryan, J. E. (2004). The perverse incentives of the No Child Left Behind Act. *New York University Law Review*, 932-989.

Sizer, T. (1984). *Horace's Compromise: The Dilemma of The American High School*, Boston, Mass.: Houghton Mifflin.

Smith, M. and O'Day, J. (1991). "Systemic school reform." In S. Fuhrman and B. Malen (Eds.), *The politics of curriculum and testing*. Philadelphia: Falmer Press.

Sonnenberg, W.C. (2004). *Federal Support for Education FY 1980 to FY 2003*. U.S. Department of Education. Washington, D.C.: National Center for Education Statistics.

Stotsky, S. and Finn, C.E. (2005). *The State of State English Standards 2005*. Washington, D.C.: The Thomas B. Fordham Foundation.

Viadero, D. (1994) "Teaching to the Test." *Education Week*, Vol. 13, Issue 39 (Extra Edition), pp. 21-25. July 13, 1994.

American Federation of Teachers | 555 New Jersey Avenue NW Washington, DC 20001

© American Federation of Teachers, AFL-CIO. All rights reserved.  
Photographs and illustrations, as well as text, cannot be used without permission from the AFT.