

1

Why Standards and Tests Are on the Front Page

ABOUT THIS CHAPTER

History tells us that the degree of public concern with education has varied through time. As far back as the Egyptians, Greeks, and Romans, and considering as well the carefully prescribed ritual training of youth in tribal societies, formal schooling has been the hallmark of stable human communities. It is a reciprocal relationship, however: Formal education connotes stability but also bears the responsibility for maintaining it. It is not surprising, then, that in relatively peaceful times positive public attention is drawn to how we educate our youth and what we teach them. Interest in education grows when there are spare energies and resources to invest. In times of stress, however, attention comes again in response to negative evaluations of the readiness of youth to protect the future. If one generation is threatened, then the next must be prepared to survive. It is the natural order of life on our planet. In order to guarantee the survival of the species, a plant compromised by drought or disease will often use its diminishing energy to produce the best blooms just before it dies.

Education is about transmitting the culture so that it may survive in the next generation. Public education in this country is also about closing cultural gaps—gaps that many fear as potential sources of human conflict. Public attention to education can be constructive, but educational processes

have a greater chance of success if they carefully balance the needs and standards of society with the individual needs of the developing individual. Educators must also recognize that society itself undergoes renewal, and it is their responsibility to respond to and engage in the ongoing process of change (see Dewey, 1916/1973).

With this very brief, longer-range perspective of the relationship between education and society in mind, I turn in this chapter to the recent history of this relationship and to the rapidly evolving events and setting of the current time. An understanding of history and setting is crucial for teachers who must deal with these many and often conflicting influences as they make day-to-day decisions about what and how to teach, and then must deal with the consequences of these decisions. I begin in time at the beginning of 2008 and look back at the previous three decades to search for an understanding of the *Zeitgeist* of this more recent and volatile episode in the history of American education.

AT THE TOP OF THE POLITICAL AGENDA: EDUCATION

In spite of the concurrent crises of an undefined war on terrorism, a growingly unpopular long-term engagement with Iraq, and a looming economic crisis, the subject of education and the reauthorization of the No Child Left Behind (NCLB) legislation of 2002 was again a significant agenda item for President George W. Bush in his State of the Union address on January 28, 2008. Education has been his continuing subject as well as that of many politicians as we proceeded through the first decade of the new millennium. On September 23, 2001, 12 days after the terrorist attacks on the World Trade Center and the Pentagon, the U.S. Congress was grappling with and making decisions on President Bush's education plan (Associated Press, 2001). Less than a month later, three days after the U.S. attack on Afghanistan, 16 state governors were meeting at an IBM center in Palisades, New York, for the fourth education summit. The location is about five minutes from where I teach, but I was not originally invited; neither was a fair representation of other teachers and principals. A working relationship with IBM, however, did allow me to attend the opening session. Michigan governor John Engler gave the rationale for their presence in these dire times, arguing that American strength can only be maintained with an educated population. "We're in a war," he said. "We want to secure ourselves from enemies internal and external. Ignorance, lack of knowledge, poorly developed skills, these are the kind of internal enemies we can do something about" (Wilson & Weiner, 2001, p. B1).

Just a month before the September 11, 2001, attack, President Bush had addressed the predominantly African American National Urban League. He described the state of many urban schools as a “great and continuing scandal” and declared, “Rarely in American history have we faced a problem so serious and destructive on which change has come so slowly” (Bruni, 2001). A major agenda item for the ensuing 2001 summit conference was the problem of the ever-widening gap in test scores between white and minority students. The solution offered by President Bush and supported by both houses of Congress was new legislation requiring annual standards-based tests for students in Grades 3–8.

President Bush was not the first to raise the issue of higher standards. The previous president, Bill Clinton, proposed big spending increases for Department of Education programs. His 1998 State of the Union speech, of which about one fourth was devoted to education, reflected his emphasis on education and his interest in national standards and assessments. Clinton recognized the need to get the support of states on this issue as he praised the state of Georgia for its academic standards and its scholarship program. He implored that the “American people respond to the challenge . . . to make American education the best in the world, to understand that it won’t be done overnight, and not to be afraid of trying to reach higher standards” (Hoff, 1997, p. 1).

WHY THE FOCUS ON STANDARDS AND TESTS?

In an interview of national leaders (Lehmann & Spring, 1996), Donald Stewart, president of the College Board, said that educational standards are necessary because we need to answer the question, “Are we better in terms of something?” (p. 3). Former assistant secretary of education and conservative school policy spokesman, Chester Finn, added that, “standards are only meaningful if you also answer the question: ‘How good is good enough?’” (p. 6). Tests with predetermined expectations or “proficiency standards” seem to answer that question. The term *standards* has now replaced the traditional educator’s term, *objectives*. Both terms essentially entail a process of coming to consensus and producing explicit statements of the elements of the American culture worthy of transmission. These statements, in essence, become the structural frame of the written curriculum. There are, in addition, many unwritten components of the curriculum and variations in organization, setting, and detail that we will discuss in the chapters that follow. As clarified statements that constitute the curriculum, standards parallel traditional goals and objectives as well as outcomes.

The term standard, however, has an implication of high levels of expectation and monitoring that were not commonly connected to the widely used educational objectives suggested by Tyler (1949). Although current standards documents include both the consensus statements of “what students should know and be able to do” and statements that provide the measures of “how well they know and are able to do,” the greatest political emphasis has been on the formal instruments, the tests that use the measures to monitor performance. It is the monitoring process on national and international levels that has energized the rising concerns for our education systems. And it is the monitoring process that has the greatest potential for both positive and negative educational consequences.

Roderick Paige, President Bush’s secretary of education in 2001, explained the significance of mandated tests as a solution to the education gap with this remark:

When states commit to using assessment data—and by this I mean breaking down results and holding schools accountable for the performance of all of their students—they see real improvement in student achievement. President Bush and Congress can lay the groundwork for reform, but state standards and assessments are the real mechanisms for improving student achievement. (U.S. Department of Education, 2001)

THE NO CHILD LEFT BEHIND LEGISLATION

The NCLB legislation, a revised version of the Elementary and Secondary Education Act (ESEA), was passed overwhelmingly by Congress and signed by President Bush in January of 2002. The law mandated his suggestion for state-developed tests in mathematics and reading at Grades 3 and 8 and high school. This was extended to all Grades 3–8 by 2005–2006 for these subjects, and a science assessment was added for Grades 3, 8, and high school beginning in 2007–2008. The original legislation also increased federal funding for 2002 to more than \$22.1 billion for America’s elementary and secondary schools, a 27% increase over 2001 and a 49% increase over 2000 levels. New specific directions for how federal dollars are spent—a form of sanction to provide accountability—included their use to cover the cost of mandated private tutoring and the transportation of students from failing to more successful schools. Many state representatives were concerned about the cost of the tests. Rhode Island’s Commissioner of Education Peter McWalters cited a cost of \$4 million and expressed concern that

the expense would divert dollars from other needs. Conference chairman and former IBM head Lou Gerstner, however, suggested an even more costly solution to the problems presented: increasing the salary of teachers (Steinberg, 2001).

As a measure of “performance standards” or “how good is good enough,” state-determined proficiency levels must be attached to each test. With the proficiency standards as a guidepost or benchmark, schools are identified as failing if they fail to improve their test scores and move closer to the standard as judged by year-to-year progress or Adequate Yearly Progress (AYP). States must develop measurable objectives for improving the performance of all students, including the economically disadvantaged and those with disabilities or limited English proficiency. The results of the tests must be disaggregated to show the performance of each of these groups separately, and 95% of each of these groups must participate in the tests. The ultimate goal was to bring every student to the proficiency level by 2013–2014.

Schools that fail to meet their AYP objectives for two consecutive years are identified as needing improvement and must provide students with the opportunity to transfer to better performing schools or support the cost of private tutoring. Table 1.1 summarizes the consequences and rewards specified in the legislation (<http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>).

The initial legislation was supported by congressional democrats and by teachers’ unions, including the national organization, the American Federation of Teachers. The law, however, became increasingly unpopular with teachers, who disliked the loss of control and emphasis on preparation for the now-crucial exams that the law brought. The prospect of teacher evaluations and salaries based on student performance was met with even greater resistance. Many Democrats, including Senator Edward M. Kennedy, complained bitterly that Mr. Bush had reneged on a promise to provide more federal aid to help low-scoring schools improve. Some states, led by Utah, sought to rebel against the law’s strictures. And with every new round of data, academics debated whether new scores showed the bill’s impact, positive or negative.

Only minor changes were proposed as Congress took up the question of the law’s reauthorization in the spring of 2007. Quick action was, however, delayed by doubts raised by Democrats, many of who represent suburban school districts that faced the prospect of penalties, and by Republicans who remained skeptical about the expanded federal role the law had brought. Nevertheless, civil rights groups and many of the law’s original supporters, like Senator Kennedy, pushed for a compromise in which tough requirements would be paired with increased aid (Gootman & Medina, 2007).

Table 1.1 Elementary and Secondary Education Act (2002) and No Child Left Behind Sanctions

- If a school fails to make Adequate Yearly Progress (AYP) for two consecutive years, it will be identified as needing improvement and must develop improvement plans incorporating strategies from scientifically based research.
- School districts will be required to offer public school choice (unless prohibited by state law) to all students in the failing school no later than the first day of the school year following identification. The district must provide transportation to the new school.
- If a school fails to make AYP for a third consecutive year, the district must continue to offer public school choice and provide Title I funds (approximately \$500 to \$1,000 per child) for low-achieving disadvantaged students in the school to obtain supplemental services—tutoring, after school services, or summer school programs—from the public- or private-sector provider selected by their parents from a state-approved list.
- Twenty percent of Title I funds at the local school district level must be used for public school choice and supplemental services.
- If a school fails to make AYP for a fourth consecutive year, it will be subject to increasingly tough corrective actions—such as replacing school staff or significantly decreasing management authority at the school level. If a school continues to fail, the school could ultimately face restructuring, which involves a fundamental change in governance, such as a state takeover or placement under private management.
- Schools that meet or exceed AYP objectives or close achievement gaps will be eligible for State Academic Achievement Awards.
(<http://www.whitehouse.gov>)

In 2002 historian Diane Ravitch (2002) expressed the need for the law: “In response to the terrorist attacks, U.S. public schools must reclaim their vital role—preparing students to become informed citizens who will preserve and protect democracy” (p. 6). Five years later in 2007, however, she expressed her disappointment with NCLB.

Despite the rosy claims of the Bush administration, the No Child Left Behind Act of 2002 is fundamentally flawed. The latest national tests, released last week, show that academic gains since 2003 have been modest, less even than those posted in the years before the law was put in place. In eighth-grade reading, there have been no gains at all since 1998.

The main goal of the law—that all children in the United States will be proficient in reading and mathematics by 2014—is simply unattainable. The primary strategy—to test all children in those subjects in grades three through eight every year—has unleashed an unhealthy obsession with standardized testing that has reduced the time available for teaching other important subjects. Furthermore, the law completely fractures the traditional limits on federal interference in the operation of local schools. (Ravitch, 2007, Op-Ed)

Prior to the NCLB legislation, some states had set proficiency standards based on the concept of “high expectations.” Teachers and administrators involved in the process of setting these did not realize that these levels would then be attached to tests that determined fund distribution and highly public consequences for failure to make progress. They also did not predict cash rewards for those whose schools did well! Experimental test-related merit pay systems for teachers and school administrators are under way in the states of Colorado, Florida, North Carolina, Iowa, and others. New York City also rewards its superintendents and school level administrators (Goodnough, 2002) and has recently suggested the less than well-received rewards for teachers.

An interesting outcome of the parameters of the legislation, which measure the year-to-year or value-added progress toward state-set proficiency levels and penalize states that do not meet the guidelines, is that states that had originally set higher levels of standards achievement as a mark of proficiency were then at a disadvantage. Individual schools that were closer to the proficiency standard in baseline measures also made less progress than those that were significantly deficient. The irony of this, in relation to the cost for NCLB sanctions, is that in a state such as New York, where low-income students scored the seventh highest in the nation on a national test, their achievement on the New York State test was less than the predetermined and relatively high state proficiency level. Federal NCLB legislation then required that 17% of the low-income students be given the option of tutoring and transportation for school transfers (Linn et al., 2002). From the perspective of rewards for performance, rather than sanctions, New York City principals, who would receive cash salary bonuses if their schools made good progress, would also be penalized if their schools started out higher up or closer to the proficiency level.

In some cases, states lowered the standards to overcome the costly requirements for failure to make progress. In Ohio, the government responded to the realization that one third of its low-income schools would require busing by lowering the state proficiency standard for schools from 70% of their students proficient to a new standard of only 42% of their students proficient. In Louisiana, students will be considered proficient when

they score at the “basic” achievement level on their state’s assessment. Connecticut schoolchildren will be deemed proficient even if they fall shy of the state’s performance goals in reading and mathematics. And Colorado students who score in the “partially proficient” level on their state test will be judged proficient (Hoff, 2002). Even Massachusetts, which has been persistent in its requirement that students pass its MCAS (Massachusetts Comprehensive Assessment System) exam before graduation, reconsidered its standards (Stein, 2002). The tight Massachusetts standards were backed up by special classes and opportunities to repeat the exam. Eighty-one percent of the seniors passed the test on the third try—up from 68% on the first (Rothstein, 2002; Vaishnav, 2002a).

Another problem with the state tests are the instruments themselves and the scoring method. Hursh (2007) reports that “almost every recent standardized exam in New York has been criticized for having poorly constructed, misleading, or erroneous questions or for using a grading scale that either over- or understates students’ learning” (p. 504). The passing rate for the exam can be increased or decreased simply by adjusting the cut score or the score which determines proficiency. This can turn a low percentage of correct answers into a pass or a high percentage of correct answers into a failure. On exams that students are likely to take as part of their graduation requirement, the state education department makes it easier for students to pass by lowering the cut score. Conversely, the exams for the advanced, non-required courses, such as physics and chemistry, have been made more difficult. Some school districts have kept students behind or purposely moved students out of school to improve their test results (Taylor, Beaudoin, & Goldschmidt, 2007). The nationwide high school graduation rate has slumped since the enactment of NCLB.

An in-depth study of the effects of NCLB on 202 high schools by Balfanz, Legters, West, and Weber (2007) revealed that the state measures of AYP were confusing and not fairly applied. For example, schools with a higher number of subgroups, which may be an indication of legitimate and worthy attempts to deal with individual needs, are more likely to not meet the AYP standard. States also have different baseline high school graduation rates. Schools in states with beginning low rates could show AYP more readily than states with higher baselines.

A solution to the problem of variations in the proficiency levels and content of individual state tests included in the NCLB legislation is the use of the formerly optional National Assessment of Educational Progress (NAEP) tests in addition to the state tests. The NAEP has been used primarily as a national sampling benchmark and sometimes as a comparison to test the validity of state tests. Under the NCLB Act, all states taking money under the federal Title I program for disadvantaged students are required to take part in

the NAEP's reading and math exams (starting in 2003). While the results did not play an official role in evaluating states' definitions of proficiency, researchers and others compared state rankings on their own tests with those of NAEP to see which states have set high goals and which ones haven't. Inevitably, the final arbiter of states' definitions has drifted toward the NAEP (Hoff, 2002). Linn and colleagues (2002), however, suggest that this, too, has problems. They remind us that the NAEP may differ considerably from state tests and the standards they measure and that the national test also has very stringent standards for proficiency. Year-to-year comparisons may in themselves be questionable, because groups of students in the same school and teaching conditions vary from year to year.

As an example of the deficiencies that might be revealed, Balfanz et al. (2007) discovered that the percentage of students in a state scoring at proficient or basic on the most recent eighth-grade NAEP mathematics exam did not affect the likelihood that a school would make AYP. When the NAEP and NCLB test performances for low-performing high schools were compared, they were about equal, but that equality did not determine whether they met the state AYP criteria.

The NCLB legislation also contains specific recommendations for the teaching of reading, as well as more money for charter schools and for training teachers. It particularly states that instruction should be research based. The problem with educational research, however, is that the many difficult-to-control variables leave some unanswered questions about what works best (see Chapter 7).

ECONOMIC INFLUENCES AND SOME WORDS OF CAUTION

Responding to critics of the move to national testing in his opening remarks to the 2001 summit conference that I attended, IBM chairman Louis Gerstner said, "But if you listen closely, what you hear is a pathetic willingness to sacrifice an entire generation, and deny them their shot at a better chance, a better future, and a better life." I am sure he was also concerned about future generations. Education is about transmitting the culture from generation to generation. Public education in this country is also about closing cultural gaps, gaps that many fear as potential sources of human conflict. But, why all the interest at this moment in time? What is the American people's attitude toward education? Why are American businessmen concerned? Are the problems as serious as the critics make them appear? Are the proposed solutions valid? What do the educators think? Can we make our schools better?

Businesses, facing a shortage of technologically proficient personnel and challenged by the need to provide training for growing numbers of employees in what they deem the basic skills, rally to support the feelings of mistrust of the public education enterprise. Politicians pick up on the issue. Governors and the president make the improvement of education a major part of their agendas. Delaine Eastin, the California state superintendent of public instruction, took her responsibility with a global perspective: “The reality is that we not only need to be able to compare across our nation—which we get in a hit-or-miss way through things like the SAT or ACT scores—but we need to be able to compare across the globe, so that we know that California’s children can compete with the German and the French and the Japanese children” (Lehmann & Spring, 1996, p. 3).

Business community leaders and governors emerged from the education summit meetings with agendas for raising and equalizing standards across the country and assuring accountability for them through federally legislated measures. This agenda was based on juxtaposed concerns about this country’s economic health, employment needs, and reports of disappointing performances of our students on state and international competency tests. In a report prepared for the National Alliance for Business, Nelson Smith (1996) quoted the cost to business for continuing and remedial education and expressed the opinion that it is “an expense that business should not have to bear” (p. 5). Examples of costs given in the report included the following: 10% of \$75 million spent by MCI for basic-skills remediation, \$700,000 spent by Polaroid for basic English and math, \$1,350 per employee annually spent by Motorola. The author also cites, as evidence of the growing need, an American Management Association report that whereas in 1989 only 4% of American businesses provided remedial training, by 1994 “the figure had jumped to 20%” (p. 5). Smith continued with references to the hidden costs to business, to the overall costs to taxpayers for remedial education, and to the ultimate costs to the public for the consequences of poor academic and readiness-to-work skills: welfare and delinquency. The remedy the report suggested for all of these ills was higher education standards and vigilance on the part of business leaders. American businesses did find their own way to overcome some of these costs. By 2008, the progress of global-based technology allowed a good portion of the business cost to be diminished by the outsourcing of jobs to the large foreign population of educated and competent individuals willing to work for smaller salaries. The resulting loss of American jobs, however, has been a significant contributor to the economic crisis.

Another report by former astronaut John Glenn, “Before It’s Too Late” (2000), underscored the poor performance of the nation’s students on the international tests and listed four reasons why students needed to be competent

in math and science: the interdependent global economy, the need for math and science in everyday decision making, national security interests, and the value of knowledge in our common life, history, and culture. Common sense tells us that this is a serious condemnation of the educational enterprise by powerful people who are not professional educators. Their focus is mainly on the measured negative results, without regard to the many unmeasured positive outcomes. After all, how did this country achieve its technological superiority? More critically, the noneducators limit their analysis of causes for problems to the lack of a common core of criteria for promotion and graduation—and then they suggest measurement of this common core with assessments that in themselves may present problems. This proposed solution to existing problems in the education of our children completely overlooks the compounding and defeating affects on the process of education created by the societal and economic problems it hopes to defeat. It neglects the realities of an issue that is of equal concern to most educators: the issue of equal opportunity to learn. Once more, it suggests that schools bear the burden for the correction of societal ills while it offers little in the way of respect for the educators who must bear this burden.

An interesting report on schools in Japan may cast a light on the danger of using the process of education to promote societal and economic interests without consideration of the periodic and sometimes drastic variations in these interests. A recent epidemic of disruptive student behavior in the traditionally effective and disciplined Japanese school systems has been related by some to a failing economy and lack of parent support for systems on which they formerly relied. Whereas in previous Japanese economic periods hard work guaranteed a job and success, in changing economic circumstances it no longer does (French, 2002). The state of the American economy is now precarious. Will it change what we think is important? Will our singularly educated population be prepared to deal with disappointments not of their own making? Will they be open-minded to alternatives?

The NCLB legislation was the essence of the government response to the predictions and concerns of business about the declining ability of the United States to compete on an international basis because of the failure of its educational system. The legislation also incorporated the business emphasis on competition, privatization and accountability (see Hursh, 2007, and Chapter 2). Sanctions included support for privately run charter schools and shifts of control from local government. In the year 2008, concerns over the economy and the failure of private health care became a campaign focus for presidential candidates. When the economy reached a crisis stage, the government quickly interceded. Much of the blame was placed on the failure of government regulation and the resulting unbridled competition of private business. The government offer of financial support for failing lending

institutions was accompanied by regulatory mandates. It may also be the time to review the methods, support, and focus of the experiment called No Child Left Behind.

THE MEDIA INFLUENCE OPINIONS ON STANDARDS

A vigilant media recognizes the sensitivity of the American public to any threat to its economic and military dominance and increases the intensity of its coverage. It casts broad-brush aspersions on conscientious teachers, who begin to feel less confident, more resentful and resistant. It makes boards of education, concerned about their constituent's support, more vulnerable to the temptation to make hasty and ungrounded decisions. What is wrong with what they are doing? Why do outsiders see national standards as a solution to the problems?

Mortimer Zuckerman (1996, p. 128), the editor-in-chief of *U.S. News and World Report*, asks if we are content as a nation to be second, third, or fourth, and if we place greater importance on the self-esteem and happiness of our children than on what they know and can do. He accepts the financial constraints of a growing population and costs but insists that we need to get better results at the same time. Zuckerman makes a very valid point about the mobility of our population and identifies higher and uniform national standards as the needed solution. "Science does not change because it is taught in Oregon or Florida." He decries the "dirty little secret" that we already have informal national standards in that we all use similar lessons and textbooks and then argues that these represent minimal competency rather than the higher levels expected by other countries. Zuckerman also presents the underlying agenda that many educators suspect in his prediction, that "higher standards are the key to inducing performance-based innovation and performance-based assessment of teachers and administrators."

The press substantiates its position with the views of both noted scholars and grassroots proponents. Diane Ravitch of the Brookings Institution is quoted in a *New York Times Magazine* article (Mosle, 1996b, p. 47): "Nations that establish national standards do so to ensure equality of education as well as higher achievement because they make explicit what they expect children to learn to insure that all children have access to the same educational opportunities." Another *New York Times Magazine* article by Mosle (1996a) tells the story of Michael Johnson, an urban principal, whose school curriculum is based on preparation for the standardized tests he sees as gatekeepers for his students. He is glad that his teachers forced him to take subjects that seemed irrelevant and believes that his kids want direction.

Mosle cites author Lisa Delpit, who in writing about African American students, says that they “need skills, not fluency. . . . I’m sick of this liberal nonsense” (1996b, p. 42). It is this notion of greater equity that originally popularized the idea of standards among minority populations and aligned them with the conservatives. But leaders among these populations are suspicious of the more subjective performance assessment measures that have now been attached to new standards (see Chapter 6). They believe that standardized tests are fairer and that minority children need the “hard skills” they measure.

The media also carefully follow the opposing reactions generated by the shifted emphasis on HSSB (high-stakes, standards-based) tests. When the Massachusetts State Board of Education declared that the state MCAS exam would be required for graduation by 2003, headlines reported the protests in May of 2000. A petition with 7,000 signatures called for the law’s repeal. Protestors included the National Association of Colored People and the American Civil Liberties Union as well as teachers, parents, and students from urban school districts (Cochran-Smith, 2000). Apparently, what the legislature saw as a way to improve the instruction of minority students by holding teachers accountable, they saw as a device that “punishes students.” Students nicknamed the MCAS as the “Massachusetts Conspiracy Against Students.”

In response, the Massachusetts Board of Education reviewed findings from a task force and ordered the creation of an appeals process so high school students who narrowly fail the state exam could still receive diplomas. Education Commissioner David Driscoll reasoned that an educationally sound and fair appeals process for the graduation requirement would be essential to determining competency with integrity (Hayward, 2001). Subsequent resolutions by the Massachusetts Association of School Committees have called for local district rights to confer diplomas, and the refusal of several districts to comply with the state regulation has resulted in a federal court case (Vaishnav, 2002b).

As Congress approached the reauthorization in 2008, a coalition of Democrat and Republican members of Congress attempted to address the complaints against the law by suburban districts, middle-class parents, states with large immigrant populations, and teachers’ unions, an important group of voters. A draft of the proposed legislation would distinguish between schools failing across the board and those where only some student groups failed to meet annual yearly progress goals. It would address the concerns of teachers by allowing states to consider more than annual math and reading scores in labeling a school as deficient. Other measures, including tests in history, science and civics; graduation rates; and Advanced Placement tests, would be added. The changes would also allow students not fluent in English to be tested in their native language for five years.

Unfortunately, a sizable Congressional opposition group and other agenda items resulted in a delay on the reauthorization bill.

Somehow, we may have missed the boat on the communication of what is good about our schools and what is needed. Can negative reports and mandated HSSB tests with questionable validity be turned into useful mechanisms for productive change? The recommendations for ameliorating the problem are themselves grounded in several powerful but as yet unproven assumptions and beliefs about education in the United States, including the belief that the future of this country's economic health depends on the improvement of student achievement, the belief that clearly stated and uniform standards will result in higher student achievement for all students in this country, and the belief that high-stakes measures will guarantee the implementation of the standards (McCaslin, 1996; Natriello, 1996).

I discuss further details of this response in the chapters ahead. Every state has standards and matching tests at benchmark grade levels. Fortunately, the states have shown respect for the creativity and commitment of the professional education community by involving them in some of the decision making of the development process. I return later to the standards and their measures, but for now I go back a little more than two decades in history to search for the origins of this intensive effort to improve schools by raising standards.

THE ORIGINS OF THE CURRENT WAVE OF PUBLIC CONCERN

Although many writers have identified the Education Commission of the States (National Commission on Excellence in Education, 1983) Nation at Risk report (often attributed to then Education Secretary Terrence Bell) as the beginning of this current wave of concern and reform, I suggest that we move back to the mid-1960s and recall that as part of the Great Society changes, Congress enacted the Elementary and Secondary Education Act (E.S.E.A.). This was the first major federal allocation of funds for the purpose of improving education. It encouraged innovation, the acquisition of new resources, and made special provisions for the disadvantaged with its Title 1 (later Chapter 1) part. Surprisingly, this infusion of federal dollars was not in reaction to any great public concern about the failures of schools. It was just the responsible effort of a government in relatively stable economic times to promote the educational process.¹

In spite of this funding and many successful (and unsuccessful) programs,² the 1970s brought us the first evidence of declines in scores on tests, such as the Scholastic Aptitude Test (SAT; now called the Scholastic

Assessment Test), and the beginning of public anxiety about education. The SAT and its accompanying achievement tests have been longtime performance standards for the American public. Because they are used by colleges as determining factors in selective admissions, they are examples of high-stakes tests, but private ones, and not derived from state or federal curriculum documents. Nevertheless, many local curriculum policies now reflect their content. This was not always true.

In 1976 I participated in a Teachers College (Columbia University) investigation of the possible causes for a decline in scores achieved by the students in two middle-class communities. We conjectured many reasons for this diminution of scores, including esoteric ones such as the birth order of the students in school at the time and clearly significant ones such as gender. My own additional hunches had to do with differences in the value placed on the test itself among the students and their teachers, parents, and peers. The results were quite interesting. Boys considered the test significantly more important than girls and their results were better. Parent and peer influence had some influence on the scores, but birth order and other contextual factors such as time spent in last-minute preparation or relaxing the night before did not have a significant effect. A major finding was that teachers made a difference. In the school district where teachers were newly motivated to place a greater emphasis on the test, there was an improvement in overall scores. There was no attempt to investigate the articulation between the school curriculums and the test; curriculum was assumed to be no different from that for the higher scoring students of previous decades. Recent decisions by some universities to eliminate the SAT as a basis for college admissions have also been based on what some see as a disjunction between the skills tested and current high school curriculum.

The 1980s brought some new ingredients to the stew over declining SAT scores. Although there was some optimism in the scores for younger students on the NAEP,³ the scores for our older students were dismal. These were the students who were entering the work force at a time when competition from other countries became a threat. Several international test reports also showed that the United States was lagging behind many countries in the performance of its students in mathematics and science (see Chapter 2).

A Nation at Risk (National Commission on Excellence in Education, 1983) also recognized the possibility of disenfranchisement for those who could not compete within our society:

Learning is the indispensable investment required for success in the “information age” we are entering. . . . The people of the United States need to know that individuals in our society who do not possess the levels of literacy and training essential to this new era

will be effectively disenfranchised, not simply from the material rewards that accompany competent performance, but also from the chance to participate fully in our national life. (p. 7)

INITIAL RESPONSES OF THE EDUCATIONAL COMMUNITY

The educational community was not unresponsive to these reports. Supported by university researchers, several school districts experimented with “effective schools,” “outcome-based education,” and “site-based management.” The effective schools movement, which began in the late 1970s,⁴ was based on limited research that identified critical factors in schools that had achieved apparent success in spite of adversity in their environments. These factors apparently made them different and more successful than schools in comparable environments that were not as successful. The research attached school characteristics such as strong leadership; parent, teacher, and student involvement; clearly defined goals and curriculum; a safe and orderly environment; and high expectations to the likelihood of greater success for all students, including those identified as “disadvantaged.” A number of schools then launched educational improvement plans using these characteristics as criteria for change.

Although the effective schools movement did not have a large-scale following, some of the criteria were adopted into other school improvement efforts. The idea of clearly articulated goals was at the heart of the outcomes-based education (OBE) reform model. The model quickly gained a strong following—including several statewide efforts—and was probably the source of the notion of national goals and standards. William Spady, who spearheaded many of the OBE efforts, writes that, “outcome based education was ushered into the 1990’s with a resounding affirmation that this is the paradigm within which true improvement of student learning will occur for all students” (Spady & Marshall, 1990, p. 4).

The pervasiveness of the outcomes-based movement and its attention to a broad base of outcomes including attitudes and values may have heralded its demise. Critics, spurred by a small but vocal group of citizens, challenged the outcomes as “too vague, non-academic, or threatening to family values” (Vinovskis, 1996, p. 68). The terminology of outcomes-based education was quickly abandoned for the new terms related to standards and performance measures. By 1993, states that had written the outcome terminology into documents issued correctives and limited their new standards to the core curriculum areas cited in the National Goals (discussed later).

A similar demise was the fate of another aspect of some of the effective schools improvement programs: the concept of site-based management. The vision that decentralization of power would be helpful in bringing about school improvements has origins in the 1960s. In the cities, poverty, crime, drugs, and family dissolution widened the gaps in an increasingly diverse population. Greater local control was a suggested solution. Large urban school districts such as New York City and Chicago, which had highly bureaucratized and appointed central boards of education, established locally elected school boards that assumed certain, but not all, decision-making authority from the central board.

The movement toward further decentralization, as a reform in school governance that would give more power to individual schools and their parents and teachers, emerged in the 1970s in a number of places, including Florida, California, and New York. Some of these programs are still in operation. Ogawa (1994) identifies a network of “policy actors” in the development of efforts to engage teachers in site-based decision making. He labels the actors working for organizations as organizational entrepreneurs. The organizations, which included the Carnegie Forum on Education and the Economy and the National Governor’s Association, were supported in their efforts by teacher organizations, and particularly by Albert Shanker, then head of the American Federation of Teachers. As we see in Chapter 2, Shanker also supported the effort to develop and monitor standards. Although Ogawa does not identify the universities as instrumental in getting the movement going, he does say that it was their stamp of approval and documentation in articles that maintained the momentum.

The consequences of the original decentralization solution were so negative that local control in Chicago was effectively abandoned by the mid 1990s. In New York City, repeated incidents of local school board corruption caused the central board of education and its chancellor administrator to remain on constant guard. Ongoing disputes between a series of incumbent chancellors and the mayors created a highly volatile and unproductive atmosphere. The present mayor (Bloomberg) finally got state approval to abandon both the central board of education and its jurisdiction over the schools, as well as the community boards. Management of New York City schools is now directly in the hands of the mayor and his own appointed chancellor and advisory board.

The failure of decentralization to effect educational improvement in large cities has been mirrored by a similar lack of significant gains from the empowerment of single schools, teachers, and parents in local school governance. The large city efforts have been thwarted by corruption in both elections and management. The local school efforts have been hampered by the

lack of expertise, resources, and time required for teachers engaged in the additional burden of managing schools and by the unwillingness of those in power to give it to others (Solomon, 1995).

GOVERNMENT CALLS FOR REFORM

Perhaps because of the failure of grassroots efforts to reform education and continuing reports of declining student achievement, in 1989 the National Governors Association held its first National Summit on Education and outlined a major role for the states in educational reform—with a special emphasis on the creation of standards. A little over a year later, Education Secretary Lamar Alexander and President George H. W. Bush (the father of President George W. Bush) announced the “America 2000: An Education Strategy” national education goals and reform strategy (U.S. Department of Education, 1991). The introduction to the announcement refers to *The Nation at Risk* report and notes that, “we haven’t turned things around in education. Almost all our education trend lines are flat. Our country is idling its engines, not knowing enough nor being able to do enough to make America all that it should be” and that, “we’re not coming close to our potential or what is needed” (U.S. Department of Education, 1991, p. 7). Legislation responding to the goals and reform strategy was signed into law by President Clinton on March 31, 1994.

The original goals state that by the year 2000

1. All children in America will start school ready to learn.
2. The high school graduation rate will increase to at least 90%.
3. American students will leave Grades 4, 8, and 12 having demonstrated competency in challenging subject matter, including English, mathematics, science, history, and geography, and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our modern economy.
4. U.S. students will be the first in the world in science and mathematics achievement.
5. Every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship.
6. Every school in America will be free of drugs and violence and will offer a disciplined environment conducive to learning. (U.S. Department of Education, 1991)

Two more goals were subsequently added:

7. The nation's teaching force will have access to programs for the continued improvement of their professional skills and the opportunity to acquire the knowledge and skills needed to instruct and prepare all American students for the next century.
8. Every school will promote partnerships that will increase parental involvement and participation in promoting the social, emotional, and academic growth of children. (U. S. Department of Education, 2001)

The goals were accompanied by suggested strategies for their accomplishment. These include "an accountability package based on 'World Class' standards for each of five core subjects" (U.S. Department of Education, 1991) and a nationwide voluntary examination system in the core subjects. Other strategies include public reporting and reward systems such as presidential citations, merit school funding, academies for school leaders and teachers, and differential pay. The document also encourages creative experiments and major commitments by business and local communities.

The goals were obviously overly optimistic for realization by 2000. Some would believe that this is because they did not have the accountability bite of mandated tests. Several initiatives were enacted and funded by Congress. Programs for the improvement of math and science education that were previously funded by the Dwight D. Eisenhower Act received new allocations. Grants in specific response to Goals 2000 (in addition to other existing program grants) were also awarded for the State and Local Systemic Improvement program, the Goals 2000 program, the Technology Challenge program, and the Schools-to-Work transition program, but each year political wrangling threatened and undermined well-intentioned plans. Previously allocated money was rescinded and new allocations were subject to constant revisions.

INITIAL RESPONSES TO THE CALLS FOR REFORM

Following the Goals 2000 announcement, many state and local efforts were initiated for the purpose of developing new curriculum and assessment standards and corresponding reform programs. In addition to individual state and local efforts, coalitions were also formed. In 1991, The New Standards project, a coalition of six large cities, 14 states, the National Center on Education and the Economy, and the Learning Research and Development

Center based at the University of Pittsburgh, embarked on a major effort to “set very high academic standards for all students, and create a system to measure their progress” (Borthwick & Nolan, 1996, p. 3). With funding from private sources and the states and school districts that form its consortium, it operated under the premise that “what gets tested gets taught.” The coalition partners believed that better performance measures will increase the possibility of accomplishing high standards, and they focused their efforts primarily on the production of these measures in the areas of English language arts, mathematics, science, and applied learning. They may have set the stage for NCLB.

One of the most successful internal attempts by the educational community to reform education was initiated by the National Council of Teachers of Mathematics (NCTM) when it published its standards in 1989 (National Council of Teachers of Mathematics, Commission on Standards for School Mathematics, 1989). The NCTM produced and disseminated the standards and several supporting materials that have been widely adopted. I discuss the influences of state policies, the NCTM, and other professional organizations in greater detail in Chapter 2.

STANDARDS IN OTHER COUNTRIES

In other countries, where the democratic ideals of private rights are not as ingrained, the imposition of national constraints and rules is less likely to be counteracted by active and passive resistance. As an example, consider a law passed by the German government that officially changed the way certain words are spelled in order to ease the difficulties that children had in using correct spelling. The law didn’t just make a recommendation; it provided for sanctions of teachers and officials who disobey. Although there was some debate on the issue (some felt it didn’t go far enough), there were some very close and firm deadlines for implementation. Newspapers, publishers, and others are gearing up for the change. Contrast this with our nation’s century-long debate over a switch to the metric system. We are now the only developed country not using it.

Borthwick and Nolan (1996) identify six qualities of national standards in countries that outperform us. These qualities are as follows: specific, public, rigorous, high stakes, inclusive, and measurable. Japan identifies precisely what the core knowledge content is to be covered—but still leaves much of the detailed instructional decisions to teachers. In an interview for the 1996 National Summit, Diane Ravitch talked about the Japanese national standards as a model of simplicity: “They are so much simpler and clearer

and more challenging and more direct than anything I have seen from American standards writers.” Albert Shanker agreed that “standards have to be doable.” California’s state superintendent of public instruction, Delaine Eastin, reported that plain language is a criterion for her state’s standards: “very useful, usable, and readable—they won’t be written in Edu-speak” (Lehmann & Spring, 1996, p. 11).

Japan’s standards may not be verbose, but the power of its government prescribed spiral curriculum is further driven by the Japanese University Entrance Examinations that are given in three stages. There are also differences in the instructional components of curriculum. Japanese teachers emphasize problem solving in math whereas American teachers and their textbooks have been more concerned with facts and procedures. The Japanese teachers are also more likely to engage students in reflective discussion, include fewer but more real-world questions that require extended answers, and provide students with manipulatives (Stevenson & Stigler, 1992). I remember being awed by a particularly clever individual manipulative package that was given to every child to keep and use as long as it was needed. I even tried to buy them for my students.

Japanese textbooks also do a better job of what Mayer, Sims, and Tajika (1995) call *cognitive modeling* than do American textbooks (discussed later). They show students how to work out problems in detail instead of devoting space to unexplained exercises involving symbol manipulation. Most Japanese children do spend much more time on homework than children in other countries and also go to private after-school coaching programs called *jukus*, which prepare them for the exams.

France’s curriculum is published and widely available. Its yearly exams are followed by much public discussion of the questions and the results. Although Sweden has tests, they are not as rigorous as the high-stakes baccalaureate of France and the *abitur* of Germany, both of which determine student access to further education. In contrast to the beliefs of many Americans, other nations have not only caught up with us in terms of secondary student retention, but they have also insisted on uniform expectations with rigorous tests, even for those not planning higher education.

In 1988, Great Britain embarked on a major endeavor to develop specific standards in every subject area and by the mid-1990s matching tests. The original opposition to the restrictive nature of the system died down somewhat—perhaps because of the vast energies required in making necessary testing adjustments. In October 2002 (British Broadcasting System, 2002) a pronouncement by exams chief Ken Boston, head of the Qualifications and Curriculum Authority (QCA), and the man then in charge of England’s exams, said the testing system needed overhauling—with

teachers being trusted to do more assessment themselves. Boston said the QCA would produce an easy-to-understand “new generic statement on standards” by the end of the month. It would work with all the exam boards on how statistics of performance from previous years should be used to set grade boundaries. These changes may generate teacher ownership and serve as an example for the United States. I address the possibilities of teacher-constructed, standards-matching assessments in Chapters 5 and 6.

Surprisingly, with the exception of the Netherlands, where there is elaborate tracking of students and a free choice school voucher system, a study by Unks (1995) reveals that the standards in Japan and the European nations are the same for all students. Even in Germany, where at a relatively early age students are tracked into three different schools with varying emphases on the pure, the practical, or the applied, the curriculum is the same. Still, as Unks notes, the curriculum in the countries studied resembles that of the United States in that it is subject centered and “the evidence is weak, suspicious, or non-existent that the study of any particular school subject (as it is usually conceived and taught) promotes outcomes such as critical thinking, creativity, citizenship or many other desirable goals” (p. 425).

As elaborated in chapters ahead, we in this country do already have a variety of public and private standard-setting, inspection, and testing systems in place, but they are inconsistent. The growing public calls for consistency generated at the federal level have been responded to, but formal enforcement mandates are another matter. If history is our teacher, then we can predict that in this country these will not come easily.

A SUMMARY: SYSTEMIC CHANGE, RESTRUCTURING, AND REFORM

In an effort to distinguish the current attempts to change educational processes from previous educational innovations, which may have been superficial and often transitory, the terms *systemic change*, *restructuring*, and *reform* are most often used—sometimes interchangeably. Smith and O’Day may have been among the first to apply and define the term systemic reform (Vinovskis, 1996). They identified three major characteristics of systemic change:

1. Curriculum frameworks that establish what students should know and be able to do.
2. State policies that would provide a coherent structure to support schools in designing effective strategies for teaching the content of the frameworks to all their students.

3. Restructured school governance systems. (Smith & O'Day, cited in Vinovskis, 1996, p. 59)

Certainly the first of the aforementioned actions has characterized the recent efforts to restructure or reform schools. Curriculum frameworks in the form of standards are omnipresent. As expected, the autonomy of the individual states has prevailed in the development of standards and in the development of matching performance measures. The resulting variations, however, have caused some confusion and difficulties in federal attempts to enforce the implementation of the standards with federally determined sanctions that provide authority. Experiments in restructured school governance systems have thus far not proven fruitful—although there are exceptions. Unfortunately, the energies and funds required for the first action, and for resolving the resulting difficulties, may have distracted us from the second action of providing a coherent structure of support for schools in their implementation.

Some of the actions taken have neglected to recognize the need for standards for opportunities to learn. There is still a lack of equity in the kind of educational environments that our students have available to them, and it seems unfair to hold everyone to the same standard if opportunities and the funds to create them are not equal. These are the issues that have been raised by educators, parents, and students. They are the issues we have to deal with. I begin to address them in Chapter 2.

In any case, more than at any other time in this nation's history, attention has been brought to bear on American education. New controls threaten the very individual and personal system of classroom instruction that is our legacy from the one-room schoolhouse of the 1800s. These controls are formidable and in response to some widespread dissatisfaction with our accomplishments. The facts on which this dissatisfaction is based may be exaggerated and unfounded—ignorant of the many difficulties faced in a changing social and economic structure. But, nevertheless, we educators will have to respond to these by knowing and understanding what they imply and by using our creative classroom decision-making power to do the most we can to help each child reach for the standards set. This is no easy task, but it is our investment in the future.

NOTES

1. Federal aid has come in the form of categorical and block grants. There is evidence that the categorical grants that are aimed at specific purposes are more effective in effecting school improvement (Kirst, 1995).

2. The Rand Corporation (Mann et al., 1975) study of the effects of these programs was one of the first to document the resistance of public schools to attempts to make serious changes.

3. The NAEP is a federally managed sampling test. It is designed to evaluate the national program and does not hold individual schools or students accountable. It is therefore not a high-stakes test. In the 1982 administration, the success rate of students on application questions in mathematics fell another 1.1% from an already low percentile of less than 50% in 1978 (Dossey, Mullis, Lindquist, & Chambers, 1988).

4. Ronald Edmonds (1983) and Lawrence Lezotte (e.g., 1981) spearheaded this effort to improve instruction by getting schools to meet the identified criteria.