

7 THE SEARCH FOR ORDER AND THE REJECTION OF CONFORMITY: STANDARDS IN AMERICAN EDUCATION

DIANE RAVITCH

To many educators, the movement for national standards and assessments seems like a remarkable innovation, a development completely unprecedented in American history. Critics, left and right, fear that it will lead to a federal takeover of curriculum; others, aware of the basic conservatism of education, declare that it cannot work and will not happen.

Whatever the truth of these assertions, the fact is that the United States has a long history of standard-setting activities, sometimes overt and purposeful, at other times implicit and haphazard. The current movement does not come from nowhere. It is grounded in a tradition of efforts to establish agreement on what American students should know and be able to do and to determine how well they have learned.

The desire to assure that all children have access to schools that offer education of similar, high quality has been a primary reason to establish standards. Over the years, standards of various kinds have evolved—sometimes purposefully, sometimes serendipitously—to foster similarity in the quality of schooling by such means as

- the use of identical or similar textbooks;
- the specification of requirements for high school graduation or college entrance;
- the use of standardized achievement tests for promotion or college admission;
- the prescription of curriculum patterns; and
- the professionalization of teacher training, with shared norms and expectations (i.e., standards).

Although educators and educational critics periodically rebel against the constraints and stifling effects of conformity, uniformity, and standardization, a great deal of time and attention have been spent by educators and legislators trying to establish shared norms, expectations, and standards

for purposes of efficiency or equality or both. This tension between the search for order and the rejection of conformity is healthy, since we value individualism and imagination far too much to embrace mechanistic prescriptions. Nonetheless, the effort to assure equal access to a good education for all students inevitably requires some form of standards.

In the earliest days of formal schooling, before the Revolution, standards existed by default, since most students learned to read from a very limited number of books, such as *The New England Primer*. For many years after the Revolution, Webster's blue-backed speller provided something akin to a national standard. Webster's spellers were used not only in school but by adults for self-education, setting clear standards for spelling and pronunciation, as Webster had hoped. And for many years, Lindley Murray's *English Reader* was the standard text in literature. As early as the 1830s, however, school reformers complained about the burgeoning numbers of textbooks in every field. Even though there was a multiplicity of titles, the content of the books was more like than unlike. And in some fields, like reading, one or two series dominated the market, like the McGuffey's reading books. When one looks at the textbooks of the nineteenth century, it is striking how similar were the history books, the readers, and others vying for market share in the same subject area. The uniformity found in the reading materials extended to classroom methods, with rare exceptions.

American schools by and large had content standards, as defined in relatively uniform classroom materials, and they even had an implicit consensus about performance standards, with a broadly shared scale that ranged from A to F, or from 100 to 60. It was not exact, but educators had a common vocabulary with which to gauge student performance.

Lacking any state or national testing systems, the only reliable standard for student performance in early America was provided by college admission requirements. Each college had its own entrance requirements, and prospective students were examined by the president of the college or members of the faculty in specific subjects. The first admission requirement to Harvard in 1642 read: "When any Scholar is able to read Tully or such like classical Latin in verse and prose, *suo (ut aiunt) Marte*, and decline perfectly the paradigms of nouns and verbes in ye Greeke tongue, then may hee bee admitted into ye College, nor shall any claime admission before such qualifications."¹ (*Suo, vestro, or nostro Marte* was a Latin proverb meaning by one's own exertions, without any help whatever.) The only significant addition to the college curriculum in colonial times was arithmetic, which appeared for the first time in the entrance requirements for Yale in 1745: "That none may expect to be admitted into this College unless upon Examination of the President and Tutors, They shall be found able Extempore to Read, Construe and Parce Tully, Virgil and the Greek

Testament; and to write True Latin in Prose and to understand the rules of Prosodia, and common Arithmetic, and Shall bring Sufficient Testimony of his Blameless and Inoffensive Life."²

As time went by, college entrance requirements became both broader and more specific. That is, they expanded during the course of the nineteenth century to include more subjects, such as algebra, geometry, English grammar, science, modern foreign languages, history, and geography, but they became more specific about the literary works that students must master before their examination. Thus, Columbia College declared in 1785: "No candidate shall be admitted into the College . . . unless he shall be able to render into English Caesar's Commentaries of the Gallic War; the four Orations of Cicero against Catiline; the four first books of Virgil's *Æneid*; and the Gospels from the Greek; and to explain the government and connections of the words, and to turn English into grammatical Latin, and shall understand the four first rules of Arithmetic, with the rule of three."³ Although there was some small variation in the Latin works that were required, the three constants in Latin were Cicero, Virgil, and Caesar. Students prepared themselves in accordance with these content standards and presented themselves for examination at the college of their choice.

In retrospect, the requirements seem fairly uniform, yet there was still enough variation in college entrance requirements to frustrate headmasters of academies and secondary schools. Consequently, in the late nineteenth century, several associations were created to promote closer relations between schools and colleges and, especially, uniformity of college admission requirements. The collaboration was good for the schools because it relieved them of the burden of preparing students for a wide variety of entrance examinations; it was also beneficial for the colleges because it enabled them to exert influence on the curriculum of the secondary schools.

As the last decade of the nineteenth century opened, many educators believed that the addition of new subjects had turned the high school curriculum into an anarchic mess. In 1892, in an effort to promote uniformity of curricular offerings in the high schools, the National Educational Association (NEA) established a panel called the Committee of Ten to make recommendations to improve the high school curriculum.⁴ The chairman of the committee was Charles W. Eliot, president of Harvard University and one of the nation's most esteemed educators of his era. The committee included William T. Harris, the U.S. commissioner of education (and former superintendent of St. Louis), four other college presidents, and three high school principals (the tenth member was a college professor).

The Committee of Ten began its work by surveying forty high schools to learn what courses they were teaching. The survey revealed that thirty-

six different subjects were offered, including five foreign languages, six mathematics courses, four science courses, and a few miscellaneous courses like stenography, penmanship, and music. To modern eyes, this menu stands in sharp contrast to the more than two thousand course titles that have been regularly reported to the U.S. Department of Education by the nation's high schools since the 1970s.

The Committee of Ten was the first national blue ribbon panel to study the curriculum of the high school; in the century that has passed since the creation of the Committee of Ten, many forests have been felled to print the reports of numberless committees, commissions, panels, task forces, and study groups on the needs or future of American education. In 1892, there was no precedent for a national body to issue recommendations to the many thousands of school districts, nor was there any mechanism to promote or require compliance. The committee had no way of knowing whether anyone would heed its proposals. Yet its report, perhaps because of the stature of its members and sponsors, as well as the novelty of the undertaking, received widespread attention and achieved some measure of influence upon the curriculums of many schools.

The Committee of Ten had to wrestle with four difficult issues: First, how to resolve the antagonism between the classical curriculum and modern academic subjects like science, history, and modern foreign languages; second, how to promote uniformity in preparation of students for college (and, conversely, how to encourage colleges to accept modern subjects as valid for college entrance); third, how to respond to demands by some educators for inclusion of practical courses like manual training; and fourth, whether the high schools should offer different curriculums for those who were college bound and those who were not.

The report of the Committee of Ten was an effort both to establish new curricular standards for high schools and to alter the admission standards for colleges and universities. And while it is true that other influences were simultaneously at work in schools and society, both complementing and subverting the proposals of the Ten, it seems clear that the report was extremely effective in changing standards at both levels of education in the years after it was issued. The committee engaged nine subject matter conferences to examine each major subject and to make recommendations on how it should be taught, how teachers should be prepared, when it should be introduced, for how many hours each week and for how many years, and so on. Each subject matter conference considered carefully the components of the course of study and how the subject should be assessed for college admission. The subjects were Latin; Greek; English; other modern languages; mathematics; physical science (physics, astronomy, and chemistry); natural history (biology, including botany, zoology, and physiol-

ogy); history, civil government, and political economy; and geography (physical geography, geology, and meteorology).

The Ten recommended that the modern academic subjects should be equal in status to the classical curriculum for purposes of college entry. The report endorsed four model curriculums, which differed mainly in the amount of time given to foreign languages: *classical* (containing three languages, including Greek and Latin); *Latin-scientific* (containing two foreign languages, one of them modern); *modern languages* (containing two modern foreign languages); and *English* (containing only one foreign language, either ancient or modern). This recommendation implied certain things that seemed radical to some educators: first, that neither Latin nor Greek was absolutely necessary for college preparation and, second, that students should be permitted to choose their course of study, so long as it included English, mathematics, history, science, and foreign language.

The Committee of Ten lined up unequivocally on the side of educational equality. Demands were already being heard in the educational press for different kinds of education for the children of workers and the children of privilege, but the Ten rejected this counsel. Instead, the committee took a firm stand against differentiation between those who planned to go to college and those who did not. All nine of the subject matter conferences, and the Committee of Ten itself, agreed "that every subject which is taught at all in a secondary school should be taught in the same way and to the same extent to every pupil so long as he pursues it, no matter what the probable destination of the pupil may be, or at what point his education is to cease." The report concluded that "the secondary schools of the United States, taken as a whole, do not exist for the purpose of preparing boys and girls for colleges." They exist to prepare young people for "the duties of life," for which the best preparation is what we would today call a liberal education.⁵

Subsequent commentators complained that the Ten ignored those who were not college bound (the overwhelming majority of children at that time), instead of recognizing that the Ten meant that all children—especially those who were not headed for college—should have the benefit of a liberal education. The idea of the Ten was radical then, as it is now, for they proposed that all children had the intellectual capacity to benefit from an education that included foreign languages, mathematics, history, mathematics, science, and English. The Ten saw this not as a college-preparatory curriculum, but as a curriculum that would prepare all children for a rich and full life, no matter what their ultimate vocation.

One immediate result of the report was the creation of the Committee on College-Entrance Requirements, established by the NEA to promote the recommendations of the Committee of Ten. This committee, far less cele-

brated than the historic Ten, worked at formulating a common framework for college preparation, consonant with the recommendations of the Ten. It recommended that high schools adopt the use of "constants" or "units" to provide a uniform measure for all courses. The committee recommended a total of ten units: four units (or years) of foreign languages, two units of mathematics, two units of English, one unit of science, and one unit of history. In a four-year program of sixteen units, this left the student free to elect six additional units. The consequence of this suggestion was that the curricular discussion—previously focused by the Ten on parallel, equivalent *courses of study*—shifted to the concept of interchangeable, equivalent *units*. This change not only advanced the notion of a standard unit of study, but eventually served to promote the principle of electives and of equivalency among all kinds of subject matter. These *units*, initially proposed by the Committee on College-Entrance Requirements, were retitled *Carnegie units* after the Carnegie Foundation for the Advancement of Teaching defined a unit as a course of five periods each week for one academic year.

Another product of the movement in the 1890s to establish uniform standards for high school graduation and college entry was the College Entrance Examination Board. Sponsored by President Eliot of Harvard, the College Board was a fulfillment, in the words of one historian, of "Eliot's two-fold vision of uniformity of standards and flexibility of programs."⁶ The purpose of the College Board was to organize a common examination system for college admission. This arrangement assured colleges that they could continue to admit whomever they chose, regardless of test scores, without relinquishing authority to any outside body.⁷ It was an ingenious solution to a vexing problem: The private sector took responsibility for creating a standard-setting process, which left high schools free to shape their curriculums as they saw fit and colleges free to admit whomever they wished.

The College Entrance Examination Board of the Middle States and Maryland held its first examination in June 1901 in nine subjects: chemistry, English, French, German, Greek, history, Latin, mathematics, and physics. The examinations were based on standards set by recognized national committees—for example, the American Philological Association, the Modern Language Association, and the American Historical Association. In the second year, new subjects were added: Spanish, botany, geography, and drawing.

The College Entrance Examination Board had to grapple with the issue of standards in the different subject areas. Initially, it assumed that it could rely on the various scholarly bodies, like the Modern Language Association, but these proved to be uninterested in secondary school teaching. So the board created its own "committee of review," with responsibility to establish requirements for the examinations in each subject. This commit-

tee regularly assembled special commissions of school and college teachers to review and revise subject standards.

Even the act of reading and grading the examinations helped to support the implementation of what might be called voluntary national standards in the different subject areas. Each year, large numbers of teachers from schools and colleges across the country would meet in New York to read the examinations. There they would talk, discuss papers, laugh about "boners," and work informally at defining performance standards for their field. In modern terms, they networked as a community of scholars and teachers. At the same time, they shaped standards and enforced them. The grading of the examinations was a professional development seminar of the highest order.

Many secondary schools prepared their students for the college entrance examinations. By reading through old examinations and by perusing the College Board's syllabus of English classics, teachers could be sure that their students were well prepared for the examinations. But these practices led to complaints about cramming and to criticism that the form of the examination tested memory power rather than the students' ability to use what they had learned. Nonetheless, the secondary schools knew what their students had to learn to prepare for college and for the entrance examinations.

The College Entrance Examination Board provided a standard for college preparation, at least for those students who wanted to apply to the selective colleges that were members of the College Board. But many educators continued to object to college domination, and in the second decade of the century, the National Education Association sponsored a Commission on the Reorganization of Secondary Education (CRSE), which established a pattern of standards that sharply diverged from the academic emphasis of the Committee of Ten and the College Board.

Unlike the Committee of Ten, which was chaired by the president of Harvard and included college presidents and secondary school principals, the CRSE was dominated by educationists. Its chair was Clarence D. Kingsley, the state high school supervisor for Massachusetts, and its members were drawn mainly from the world of professional education and colleges of education. Unlike the conferences of the Committee of Ten, which focused on academic subjects, the CRSE established committees not only for academic subjects but also for industrial arts, household arts, vocational guidance, agriculture, and other nonacademic areas. The report of this commission, published in 1918, identified "the main objectives of education," as follows: "1. Health. 2. Command of fundamental processes. 3. Worthy home-membership. 4. Vocation. 5. Citizenship. 6. Worthy use of leisure. 7. Ethical character."⁸ Every academic subject was required to demonstrate its value in achieving these objectives; the emphasis was on

utility and social efficiency. Several academic subjects, especially the classical languages and history, were difficult to justify or rationalize in terms of the seven cardinal principles. Neither history nor geography survived as a subject; both were submerged into the new field of the "social studies." The new standard for high schools, then, was not to be based on the intellectual development of all youngsters, nor on a commitment to the ideal of liberal learning, but on preparing youngsters for present and future social roles. The goal—the standard—was social efficiency. The report endorsed differentiated curriculums, including agricultural, business, clerical, industrial, fine arts, and household arts curriculums. At the time, this seemed an appropriate approach for schools that were suddenly overwhelmed with large numbers of immigrant children, many of whom spoke little or no English.

The report gave strong support to the development of comprehensive high schools. Unfortunately, it also supported the practice of curricular tracking, whereby school officials could guide students into appropriate curricular experiences based on predictions about students' future vocation. These guesses tended to encourage differentiation based on social class, race, and ethnicity. Thus, the academic track increasingly became a preserve for the minority who were college bound—the bright, the ambitious, the children of the educated—while the other tracks were occupied by the majority of students who were directed into programs that presumably would prepare them for their future vocational roles. From the point of view of the commission and many other educators, college-preparatory programs were a waste of time for these children.

The report of the CRSE was, of course, a reflection of widely shared views in the education profession and in society. The term *academic* was increasingly used in a derisory fashion to refer to studies that had no practical value. Many educators, for many different reasons, converged on certain themes that appeared in the report: that an academic or liberal education was not appropriate for everyone; that the purposes of education should be derived from the activities of life, rather than from book learning; that book learning was sterile, "academic"; that children had different needs and should therefore have educations that were fitted to their needs; that the role of the school should change to fit the needs of society (and the needs of society were usually in the eye of the beholder). The overriding philosophy was social efficiency, and it supported curricular tracking and a devaluation of liberal education.

The two reports, separated by twenty-five years, both claimed to be based on the principles of a democratic society. The Committee of Ten believed that all children should have the experience of a common academic curriculum, that all should be educated in the same way regardless of who their parents were or what their intended destination in life; the

authors of the *Cardinal Principles* believed that the curriculum should be tailored and differentiated to meet the needs of society and of children. Beginning with similar premises, the two ended up in very different places. The compromise that was struck over time between the conflicting ideals was that the principles of the Committee of Ten applied to the academic track and the principles of the CRSE governed the vocational and general track. Or, put another way, the CRSE won. The ideas of the CRSE won out not because they were argued more persuasively but because they provided a good fit for the thorny problems that faced the schools: the problems of mass education and the problems of educating large numbers of children from poor and non-English-speaking backgrounds. Parents of these children did not demand that they be excused from the academic curriculum; what evidence exists suggests that immigrant parents did not want their children to have a curriculum different from that in the best schools. It would have been possible to educate poor and immigrant children in the way prescribed by the Ten, but to do so would have required enormous intellectual energy in the reformulation of subject matter. It was easier to teach academic subjects to academically able children than to rethink and redesign what was taught in school so that all children could learn and understand the material in the college track. To do so would have been very difficult and very inefficient in an age when efficiency was highly valued. The path of least resistance was chosen, and, in time, educators and parents came to believe that certain studies—like advanced courses in mathematics and science and history—were only for the college-bound students, a distinct minority.

At the same time that the CRSE provided the rationale for differentiation of curriculum, the availability of the new techniques of standardized testing facilitated differentiation. After the debut of intelligence tests during World War I, the testing industry mushroomed, providing a broad array of tests of ability, tests of intelligence, and tests of educational achievement.

The question of standards, as applied to the academic curriculum, was defined during most of the twentieth century by college admission requirements. These consisted of the colleges' entrance requirements—a specification of so many years or units of certain academic subjects—and the college entrance examinations. Until the development of the American College Testing (ACT) program in 1959, the only national examination for the college bound was that of the College Entrance Examination Board. The College Board, while vigilant in protecting the quality of its examinations, was sensitive to constant criticism that the examinations exerted too much influence on the high school curriculum. In the years immediately after World War I, the board took interest in the new movement for intelligence testing. Unlike the traditional college entrance examinations,

which tested what students knew, the intelligence tests claimed to test students' innate intelligence or what students were capable of doing. In 1922, the College Board expressed "favorable interest" in the use of a "general intelligence examination" and its readiness to administer such examinations when practicable.⁹ Since none of the members of the board was an expert on psychological testing, they appointed an advisory commission of experts, including Carl C. Brigham of Princeton, Henry T. Moore of Dartmouth, and Robert M. Yerkes of Yale. Brigham and Yerkes were, of course, pioneers in the development of intelligence testing. Brigham soon emerged as the driving force in a successful effort to reorient the college entrance examinations and turn them into tests of intelligence or aptitude.

Brigham and his team of psychological experts built upon the College Board's interest in determining a student's readiness for college-level work and used it as the foundation for a new test altogether, the Scholastic Aptitude Test (SAT). There were skeptics on the College Board and among the secondary school representatives, but the experts had a mantle of authority—the authority of science—that swept along the doubters.

The first Scholastic Aptitude Test was given in 1926 to 8,040 candidates, most of whom were applying to Ivy League colleges. The father of the SAT, Carl Brigham, became a salaried member of the College Board staff in 1930, where he continued to refine the SAT. Although the SAT did not take the place of the traditional written entrance examination until the Second World War, the College Board recognized its value immediately. The appeal of the SAT grew apace with faith in social science. The standardized, multiple choice test saved the College Board from the wrath of those who complained that the college entrance examinations had too much influence over the curriculum of the secondary schools. It was the perfect answer to the angry headmaster who supposedly muttered in 1902, "I'll be damned if any Board down in New York City, with a college professor at its head, is going to tell me and my faculty what or how to teach!"¹⁰ The SAT tested linguistic and mathematical power and had no connection to any particular curriculum. This left secondary schools free to require whatever they chose. The literature curriculum, which had been anchored by the college entrance examinations for many years, was completely abandoned by the SAT, allowing secondary schools to teach whatever books they wished and, if they chose, to drop the traditional classics altogether.

During the 1930s credibility shifted from the old-style written tests to the new-style SAT. Although some of those on the College Board must have been uneasy about the shift to intelligence / aptitude testing in place of achievement testing, Brigham became "the policy maker" of the board, "and whatever route was taken for the next few years was charted by him."¹¹ Colleges began to recognize the practical value of the SAT for

predicting students' ability to do college-level work, and each year the number who took the SAT increased. Then, on Pearl Harbor Day, 1941, at a meeting of representatives of Harvard, Yale, and Princeton, a decision was made to cancel the June essay examinations because of the war. Normally there would have been an outcry from conservative educators, who continued to believe that an essay examination was preferable to a multiple-choice test of aptitude. But the war emergency hastened what at the time seemed inevitable. With this decision, the SAT became *the* college entrance examination for the nation's most prestigious colleges and universities. The advent of the machine-scored examination meant, first, that the standard-setting force of the traditional (i.e., written) college entrance examination was eliminated and, second, that the standard-implementing activities of those teachers and scholars who met annually to read the examinations were terminated.

During most of the 1940s and 1950s, the College Board claimed that the purpose of the SAT was not to influence standards but to help colleges identify students who were ready for collegiate studies. However, other examinations offered by the College Board defined and supported educational standards, such as advanced placement (AP) examinations and achievement tests. Over time, it became clear that the AP exams and the achievement tests bolstered high school academic standards at the same time that the SAT undermined them. With its focus solely on mathematics and verbal skills, the SAT was virtually curriculum-free. By implying that college admission would be decided by *aptitude* rather than *achievement*, the SAT encouraged the widespread belief that innate ability, not effort, was what really counted. Although the Educational Testing Service contended for many years that scores on the SAT were not affected by coaching, students did improve their performance by attending special coaching sessions or by learning test-taking skills. Until the latter 1960s, the SAT was buttressed by high school graduation requirements and college entrance requirements. When the latter two were reduced in the late 1960s and early 1970s (in response to student demands), the SAT was left standing almost alone as the guardian of standards, a role for which it was not designed and to which it was not equal.¹²

College entrance examinations were one source of educational standards; another was mandated testing, which was introduced on a broad scale in the early twentieth century. In the 1890s, educational reformer Joseph Mayer Rice introduced what may have been the first test to be administered to a large national sample (in spelling). Edward Thorndike of Columbia University introduced many standardized achievement tests. Thorndike's colleagues at Teachers College helped to spread standardized achievement testing as part of the school survey movement, which assessed the quality of numerous school districts around the nation and

reached its peak in the second decade of the twentieth century. Many schools were administering standardized achievement tests in spelling, arithmetic, reading, and other subjects before the First World War. The use of widescale intelligence testing during the war, coupled with the general admiration for science and social science, helped to popularize standardized testing in the 1920s. During that decade, hundreds of intelligence tests and achievement tests were produced for use in the schools. Both kinds of tests were used to sort children according to their ability and to place them into appropriate programs.¹³

Since the 1960s, the volume of criticism of all kinds of standardized testing has escalated rapidly. Two points should be made about the general use of standardized testing of both intelligence and achievement. First, both were used to allocate students to different kinds of educational opportunities, in keeping with the curricular differentiation advocated by the CRSE. Second, the widespread adoption of standardized achievement tests relieved states and districts of the necessity of consciously setting their own academic standards. The critical point is that educators relegated the all-important task of deciding what children should know and be able to do to the commercial testmakers.

A similar story can be told about the role of textbooks as a standardizing element in American education. Woodward, Elliott, and Nagel estimated "that from 75% to 90% of classroom instructional time is structured by textbook programs."¹⁴ Produced for mass market sales in a highly competitive marketplace, textbooks are written to satisfy the largest buyers, especially the textbook-adoption committees in large states like Texas and California. Because the textbooks have such an important role in determining what content is taught and because they are so widely used as a basic instructional tool, they in effect determine what children learn. How do textbook writers decide what children should know? We hope that the writers include among them some scholars and teachers; the writers review the contents of the leading textbooks in the field and the requirements of the textbook-adoption committees in a score or so of states. Through this serendipitous process, the writers arrive at content standards that provide the framework for their product.

With the passage of the Elementary and Secondary Education Act in 1965, standardized achievement testing became securely entrenched because the law required regular testing in schools that receive federal funding through Title I or Chapter I. So the tests, always important, gained an institutional foothold and became a mandated element in the programs of a large proportion of public schools. New programs of statewide testing were introduced in the 1970s to meet the requirements of the minimum competency movement, which demanded proof that students had achieved basic skills before graduating or being promoted. During the

same era, textbooks became more uniform than ever, as big companies gobbled up little companies and as a small number of textbooks in each field captured a large percentage of the market. By 1987-88, forty-five states and the District of Columbia were using some kind of statewide test; twenty-five of those entities were employing a commercially developed, nationally normed test.¹⁵

The proliferation of mass-market textbooks and standardized tests has been criticized by many educators on a wide variety of grounds, but both texts and tests continue to be widely used and enormously influential on what is taught in the classroom. The relevant point, for this narrative, is that both texts and tests are prepared on the basis of assumptions about what is worth knowing. Sometimes these assumptions are explicitly and thoughtfully arrived at, and sometimes they are derived from market research (what sells) without a lot of time devoted to debating or studying what knowledge is of most worth.

The standards that are most widely shared in American education are those that have been embedded somehow in the commercial textbooks and tests, as well as in the college entrance examinations (the SAT and the ACT). Actually, we have multiple standards, depending on the curriculum track in which students are enrolled, whether they plan to go to college, and whether they have applied to a selective college or university. Students are educated to very high standards if they expect to take advanced placement examinations and if they intend to apply to one of a relatively small number of selective colleges. Students who are not planning to go to college are probably exposed to standards that are geared to minimum competency, which they may demonstrate by checking off the correct box on a test of basic skills and by tests of simple recall. Students who intend to go to an unselective college or university, one that takes almost all applicants, will encounter standards that are not demanding; the students know, and their teachers know, that they will be accepted into the college of their choice regardless of how much or how little they work in school and regardless of how much or how little they have learned in school.

The state of educational standards became a national issue in 1975, when the College Board called attention to the fact that SAT scores had fallen steadily and sharply since 1963. A study panel appointed by the College Board concluded two years later that the score declines had first been caused by the diversification of the pool of test takers but had then been accelerated by changes in the practices of the schools, such as dilution of the academic curriculum, lower enrollments in advanced courses, social promotion, less assignment of homework, and grade inflation. The panel made clear that high schools had lowered expectations for almost all students in response to what was perceived as the needs of a more diverse student body.¹⁶

It was not only SAT scores that declined during the late 1960s and most of the 1970s but also scores on virtually every other standardized test in the nation. Some states began to take stock of the quality of education, and none was as active as were the fifteen states that formed the Southern Regional Education Board (SREB). In 1981, a special task force of the SREB issued a statement called *A Need for Quality*, which called for higher standards for teachers and students, as well as higher teachers' salaries. Critical of the climate of low expectations in education, the SREB called for strengthening of the high school curriculum and of college entrance requirements.¹⁷

In 1983, *A Nation at Risk*, the report of the National Commission on Excellence in Education, appeared. That report warned of "a rising tide of mediocrity that threatens our very future as a Nation and a people." It described numerous indicators of low educational achievement and criticized specific failings in the school, especially "a cafeteria-style curriculum," low expectations, low standards, and a widespread lack of seriousness of educational purpose. Its primary recommendation was that "state and local high school graduation requirements be strengthened and that, at a minimum, all students seeking a diploma be required to lay the foundations in the Five New Basics by taking the following curriculum during their 4 years of high school: (a) 4 years of English; (b) 3 years of mathematics; (c) 3 years of science; (d) 3 years of social studies; (e) one-half year of computer science. For the college-bound, 2 years of foreign language in high school are strongly recommended in addition to those taken earlier."¹⁸

The publication of *A Nation at Risk* in 1983 was followed by a frenzy of public concern about the quality of education. Many states established task forces, commissions, study groups, and the like. Many raised their graduation requirements to demonstrate that students were expected to study more academic courses. Although it became fashionable to say that the reforms of this period achieved little, this is not accurate. Student enrollments increased in core academic subjects like mathematics and science, and course taking is directly related to achievement.¹⁹

And something else happened. A new breed of reformer emerged, as well as a new understanding of what was needed to improve student achievement. In 1983, California elected a new state superintendent, Bill Honig, who pledged to raise standards and improve the quality of public education. While tirelessly advocating increased funding for public schools, Honig started a process of reviewing and reconstructing the curriculum in each subject field. He recognized that the starting point for reform is agreement on what to teach, and he brought together panels of teachers and scholars to rewrite the state's curriculum frameworks, which serve as guidelines for textbook publishers. Since California has 11 percent

of the nation's schoolchildren and since the state adopts textbooks, the frameworks offered the state a way to influence the state (and national) textbook market. Honig also revised state testing, encouraging the development of assessments that were based on the state's curriculum rather than on national norms. In effect, Honig shaped a coherent program of systemic reform by setting new curriculum standards, procuring a new generation of textbooks and technology, and developing new tests.

The southern states had already anticipated the prescription of *A Nation at Risk* through the work of the SREB. Several governors, notably Lamar Alexander of Tennessee, Richard Riley of South Carolina, and Bill Clinton of Arkansas, took the lead in shaping reform programs that emphasized the establishment of educational standards and assessments. The organizing principle of this new state-based movement was that the focus of reform should be on results, rather than on process or inputs, and that states should press for higher educational achievement while reducing regulations and other procedural burdens on schools. In 1986, Governors Alexander, Clinton, and Riley worked together on a report for the National Governors' Association, titled *Time for Results*; on behalf of the other governors, Lamar Alexander wrote that "the Governors are ready for some old-fashioned horse-trading. We'll regulate less, if schools and school districts will produce better results."²⁰ Thus, the governors most actively involved in state-level reforms moved to identify standards, by which they meant *what students should know and be able to do*, and to develop assessments to provide information about whether students were making progress toward the standards.

Meanwhile, there was increasing demand by policy makers for information about student achievement. In 1986, eight southern states (Arkansas, Florida, Louisiana, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia) administered the tests of the National Assessment of Educational Progress to a representative sample of students. They did so to obtain "the most current and reliable measure of how their students' achievement in reading and/or writing compares to truly national and regional results." They also wanted to establish benchmarks "to gauge their students' relative achievement levels." Winifred L. Godwin, the president of the Southern Regional Education Board, explained the southern states' interest in assessment: "The obvious question is—How will we know that we are making progress? Many measures of progress will be important, but none will surpass student achievement."²¹

In 1987, a study group led by Lamar Alexander and H. Thomas James proposed an expansion of the National Assessment of Educational Progress to include state-by-state comparisons (Hillary Rodham Clinton was a member of the study group). Explicit in this recommendation was recognition that "state and local school administrators are encountering a rising

public demand for thorough information on the quality of their schools, allowing comparison with data from other states and districts."²² Implicit in this recommendation was the suggestion that school districts ought to be allowed to participate directly in the National Assessment to learn how their students were doing compared to other districts. Congress authorized trial state assessments for 1990, 1992, and 1994 but showed little interest in expanding the assessment to the district level. Even the state-by-state assessment was controversial, since many educators saw no value in the comparisons and recoiled against the drift toward a national test.

The state-level activity by policy makers and elected officials reflected a somewhat commonsense understanding that the effort to improve education must begin with an agreement about what children are expected to learn, that is, content standards. Traditionally, this agreement had been expressed in Carnegie units, defined by states as years of science or mathematics or English needed to graduate. But Honig in California and the southern education reformers moved beyond Carnegie units to seek greater specificity, to identify what children were expected to know and be able to do. And wherever there was standard setting there was also new interest in finding some reliable means of measuring student progress toward meeting the content standards, thus increasing the search for a test or an assessment that would permit comparisons across states, districts, even schools.

As many states wrestled with the difficult process of defining their standards, the National Council of Teachers of Mathematics (NCTM) embarked on a course that would alter the national debate about standards. In response to the criticisms expressed in *A Nation at Risk* and other national reports about the state of mathematics and science education, the NCTM decided to develop a new K-12 mathematics curriculum.²³ In 1986, the NCTM created the Commission on Standards for School Mathematics, which led a broad consensus process involving large numbers of teachers, supervisors, mathematics educators, and mathematicians. Writing teams met during the summer of 1987; they reviewed state curriculum frameworks and the curriculum standards of other nations. Ten thousand copies of the draft document prepared that summer were circulated and reviewed. The writing teams revised the draft standards in 1988, and the NCTM standards were published in 1989. The chair of the NCTM Commission on Standards, Thomas Romberg of the University of Wisconsin, wrote that the writing teams were given the following charge:

1. To create a coherent vision regarding:
 - a. what it means to do mathematics;
 - b. what students need to do in learning mathematics;

- c. what teachers should do in teaching mathematics;
 - d. what the emphasis of the curriculum should be;
 - e. what it means to be mathematically literate in a world that relies on calculators and computers to carry out mathematical procedures, and, in a world where mathematics is rapidly growing and being extensively applied in many fields.
2. To create a series of standards for the curriculum and for evaluation that articulates this vision.²⁴

The NCTM gave three reasons for developing standards: first, to ensure quality (i.e., to "ensure that the public is protected from shoddy products"); second, to express expectations (just as, for example, the American Psychological Association sets standards for tests to establish their reliability and validity); and, third, to establish "criteria for excellence." This last reason meant that the NCTM wanted to replace low minimum standards with a goal that would become a stimulus for change; as Romberg explained it, the standards would be "the 'flag' around which teachers can rally for support" and would serve as "an informed vision of what should be done, given current knowledge and experience."²⁵

The central concept of the NCTM standards is that "knowing" mathematics is "doing" mathematics. The emphasis throughout is on active learning, problem solving, reasoning about mathematics, and communicating mathematically. The standards are intended not just for the college bound, but for all students. Their goal is to transform mathematics education so that all students are able to develop mathematical power and to apply mathematical thinking. The standards are intended to replace rote memorization and drill with thinking, estimating, questioning, and figuring things out, and they provide teachers with numerous examples of suggested classroom strategies rather than with a detailed list of requirements.

The NCTM standards had their detractors, to be sure; there continued to be educators who preferred drill and practice and believed that the NCTM standards veered too sharply away from computation. And there were no guarantees that the standards would produce higher achievement because they were untested. But the NCTM standards nonetheless had a dramatic influence on the field of mathematics. First of all, they emerged from a successful consensus process: They represented the work of the nation's leading mathematics educators. Second, as standards, they were dynamic in their reach: The full implementation of the standards required the revision of instruction, teacher education, professional development, textbooks, technology, and assessment. Within two to three years after the NCTM standards were released, significant changes could be seen in every

one of these areas, as textbook publishers, schools of education, technology developers, and test makers all claimed that they were working to conform to the NCTM standards.

At the same time that the NCTM was drafting and revising and publishing its standards, education moved to the front burner as a national issue. In the fall of 1988, soon after the presidential election, President George Bush invited the nation's governors to a summit in Charlottesville, Virginia. At that meeting, the president and the governors agreed that the nation should have national goals for education in the year 2000. These goals were forged through intensive negotiations between the White House and the National Governors' Association.

In 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 per cent.
3. American students will leave grades four, eight, and twelve 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 first in the world in science and mathematics achievement.
5. Every adult American will be literate and 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 safe, disciplined environment conducive to learning.²⁶

To monitor the nation's progress in moving toward the goals, the National Education Goals Panel was established, consisting of governors and representatives of the White House and the Department of Education. (The initial exclusion of members of Congress from the Goals Panel alienated the legislative branch and made many members unwilling to support the work of the panel.) Goal 3 and goal 4 clearly had implications for the promotion of standards. For American students to demonstrate "competency in challenging subject matter," it would be necessary to define what kind of subject matter was "challenging" and what constitutes "competency."

To advance the new agenda based on the national education goals, President Bush appointed Lamar Alexander as his secretary of education in the spring of 1991. A former governor, Alexander was committed to collaboration with the governors on behalf of the National Education Goals. He shaped the America 2000 plan, which was not a new federal program but rather was a nationwide strategy to reach the goals. It involved, first, encouraging thousands of communities to develop their own local plans to reach the goals; second, promoting the idea of "new schools," schools that would "break the mold" by using time, technology, and resources differently and more effectively than traditional schools; third, supporting the development of national standards and a voluntary national achievement test; and, fourth, advocating means-tested vouchers to allow parents to choose among public, private, and religious schools.

Shortly after Alexander took office, he persuaded Congress to authorize a bipartisan commission called the National Council on Education Standards and Testing. The co-chairs of this council were Governor Carroll Campbell of South Carolina (a Republican) and Governor Roy Romer of Colorado (a Democrat), who also were leaders of the National Education Goals Panel. The thirty-two members of the council included members of the administration, Congress, and educators. After six months of intense deliberation, the council concluded that "high national standards tied to assessments are desirable. In the absence of well-defined and demanding standards, education in the United States has gravitated toward *de facto* national minimum expectations." The council stipulated that national standards should be voluntary, not mandated by the federal government; national, not federal; geared to high expectations, not minimal competency; and designed to provide focus and direction rather than a national curriculum.²⁷

To support the recommendations of the council, the U.S. Department of Education made grants to independent scholarly and professional organizations to develop content standards in science, history, the arts, civics, geography, English, and foreign languages. The grants for science and history were made even as the council was deliberating; the other grants followed the release of the council's report in January 1992. The arts, civics, and foreign languages were not explicitly mentioned in the goals, but the secretary decided that the specification of five academic subjects was not intended to exclude other important academic subjects, especially the arts, civics, and foreign languages.²⁸ The NCTM standards served as an explicit model, demonstrating the importance of a broad and inclusive consensus process and the power of standards to promote reform in textbooks, tests, instruction, and teacher education. In addition, the department supported

grant competitions to enable states to create new curriculum frameworks in the same subjects in which national standards were being developed, in effect, inviting a synergy between national and state standard setting.

By 1993, some \$10 million of federal funds had been committed to the development of national content standards in key subject areas. With additional federal funds allocated to state curriculum development not only by the U.S. Department of Education, but also—in mathematics and science—by the National Science Foundation, both money and political momentum supported widespread activity to define the content standards and performance standards at which teachers and students were to aim.

Shortly after taking office in 1993, the Clinton administration signaled its support for the standards-based reform agenda by introducing Goals 2000 to encourage the development of content standards, performance standards, and “opportunity-to-learn” standards (i.e., measures of the conditions of teaching and learning). This legislation was passed in 1994. It creates a federal agency (the National Education Standards and Improvement Council) to certify national and state standards. In the future, standards will be both national *and* federal.

Much remains to be decided: what kind of agency, if any, will oversee the development of content standards? How will these standards be revised? Will there be assessments based on the standards? Who will prepare them? Will there be any consequences (for students) attached to the results of such assessments? Can the standards be kept free of political influence? Will the standards represent criteria for excellence, or will they be diluted to a low minimum of competence?

My own view is that the purposeful effort to construct national standards is a promising undertaking that offers the hope of promoting change in many parts of the educational system. It will be a magnet for criticism, not only from those who fear the heavy hand of government intrusion, but also from educationists who distrust any emphasis on disciplinary knowledge and who find it hard to believe that children from disadvantaged backgrounds can respond to intellectual challenge.

The promise is that we as a nation can develop a clear and fruitful consensus about what we want children to know and be able to do; that this consensus will prove helpful to students and teachers and will provide the grounds for improving teacher education, assessment, textbooks, staff development, and classroom technology. The implications for assessment are obvious: Tests should be based on what students have learned, not on their aptitude or native ability. The syllabus for examinations should be made public, so that teachers and pupils know what is expected and can study what is important. It is appropriate to “teach to the test” if the test is valid, reliable, and geared to knowledge and skills that are important. The use of curriculum-free tests for important decisions—e.g., college admis-

sion and employment—serves only to certify the unimportance and irrelevance of what youngsters study in school and of the effort that they apply to their studies. With national standards, it becomes possible to base educational tests on what children learn, instead of separating what is learned and what is tested.

It seems eminently useful to describe in plain language and with forceful example what we want children to learn and to construct educational reforms around that agreement. The idea of content standards has powerful implications for instruction and assessment. It does not mean that education will be standardized, but that the expectations that we have for children will be both higher and more transparent. The act of explaining what is needed for success is an essential component of equity; other nations establish national standards both to provide equal opportunity and to encourage higher achievement. In the absence of clear expectations for all children, curricular differentiation favors the advantaged.

We learn from history that American education does have standards and that these standards have emerged in a patchwork, higgledy-piggledy, uncoordinated manner. For much of our history, our curriculum standards have been created—almost accidentally—by those who develop commercially produced textbooks and tests. What children should know and be able to do has been decided offhandedly through a sort of consensus process that occurred at the intersection between state departments of education and publishers hoping to market their products. At the high end, performance standards were established by college admission tests and advanced placement tests; for most students, however, the minimal expectations embedded in nationally normed tests became *de facto* national standards.

The challenge before us as a nation is to develop a thoughtful process to decide what knowledge is of most worth and what knowledge is most valuable to children who will live and work in the twenty-first century. History tells us that it will not be easy to do this; in fact, we know already that the fractious politics of curriculum making guarantees controversy at almost every step of the journey. Partisans with a mission will seek centralized control, if they think they can get it, to carry their message into every schoolroom; others, fearful of centralization and loss of autonomy, will resist any coordinated effort to develop content standards. But again, the message of history is that autonomy is an illusion; standards are already in place, an accidental product of decisions made for various reasons. (Was the state-adopted textbook selected because of its educative power or because it had a strong binding, good paper stock, and plentiful graphics and “mentioned” the right list of topics and names?) Could we do better as a society if we consciously and thoughtfully decided what we want children to learn and if we purposefully redesigned the customary means of assess-

ing whether and how well students have learned what was taught? Would more children achieve at higher levels if we were explicit about what was needed for success in school? Could we serve the ends of both excellence and equity by making expectations clearer to everyone involved in the educational process? There is also the substantial risk that a federal agency, governed by a politically appointed board, will be unable to endorse anything other than lowest-common-denominator standards.

The risk—and it is real—is that the effort will prove impossible because it is too complex, too controversial, too easily misinterpreted, and too radical a change. The critics are many; for most of this century, many American educators have been suspicious, often even hostile, to reforms grounded in subject matter, in decisions about what children should learn. A persistent strain in American education reacts negatively to any emphasis on knowledge, as though problem-solving skills, creativity, and other desirable educational goals were inconsistent with knowledge.

My training as a historian warns me not to expect too much, cautions me that educators have an unfortunate habit of making the best the enemy of the good, thus beating back any proposed reform that does not promise to solve all problems simultaneously or to lift all boats equally and at the same pace. Yet my participation in the process as a reformer keeps me hopeful that we will somehow muddle through, that the naysayers will give change a chance, and that we will aim for and continually pursue a synthesis of our twin ideals of excellence and equity.

NOTES

1. Edwin Cornelius Broome, *A Historical and Critical Discussion of College Admission Requirements* (New York: Macmillan, 1903), 18.

2. *Ibid.*, 30.

3. *Ibid.*, 34.

4. The National Educational Association was a forerunner to today's National Education Association. At that time, however, it was not a teachers' union; most of its members were school superintendents.

5. The report of the Committee of Ten is reprinted in Theodore Sizer, *Secondary Schools at the Turn of the Century* (New Haven: Yale University Press, 1964), or see National Educational Association, *Report of the Committee on Secondary School Studies Appointed at the Meeting of the National Educational Association, July 9, 1892; with the Reports of the Conferences Arranged by This Committee, and Held December 28–30, 1892*, Bureau of Education, document 205 (Washington, 1893).

6. Edward A. Krug, *The Shaping of the American High School, 1880–1920* (Madison: University of Wisconsin Press, 1964), 147.

7. Until the mid-twentieth century, many selective colleges routinely limited the number of Jewish and Catholic students by quotas, and they would have been

reluctant to embrace any scheme that would cause them to accept students strictly on the basis of test scores.

8. Commission on the Reorganization of Secondary Education, *Cardinal Principles of Secondary Education*, bulletin 35 (Washington, D.C.: Bureau of Education, Department of the Interior, Government Printing Office, 1918).

9. Claude M. Fuess, *The College Board: Its First Fifty Years* (New York: College Entrance Examination Board, 1967), 104.

10. *Ibid.*, 57.

11. *Ibid.*, 113.

12. Clifford Adelman, "Devaluation, Diffusion and the College Connection: A Study of High School Transcripts, 1964–1981" (Washington, D.C.: U.S. Department of Education, 1982).

13. Paul Davis Chapman, "Schools as Sorters: Lewis M. Terman and the Intelligence Testing Movement, 1890–1930" (Ph.D. diss., Stanford University, 1979), 165–80.

14. Arthur Woodward, David L. Elliott, and Kathleen Carter Nagel, *Textbooks in School and Society* (New York: Garland Publishing Co., 1988), 7.

15. Daniel Marks, "Statewide Achievement Testing: A Brief History," *Education- al Research Quarterly* 13, no. 3 (1989): 40–41.

16. College Board, *On Further Examination* (New York: College Board, 1977).

17. Southern Regional Education Board, *The Need for Quality* (Atlanta: Southern Regional Education Board, 1981), 16–20.

18. National Commission on Excellence in Education, *A Nation at Risk* (Washington, D.C.: Government Printing Office, 1983), 5, 8–9, 18.

19. Rolf K. Blank and Doreen Gruebel, *State Indicators of Science and Mathematics Education, 1993* (Washington, D.C.: Council of Chief State School Officers, 1993), 25–26.

20. National Governors' Association, *Time for Results* (Washington, D.C.: National Governors' Association, 1986), 3.

21. Southern Regional Education Board, *Measuring Student Achievement: Comparable Test Results for Participating SREB States, the Region, and the Nation* (Atlanta: Southern Regional Education Board, 1986), iii, 2.

22. *The Nation's Report Card: Improving the Assessment of Student Achievement*, Report of the Study Group, Lamar Alexander, Chairman, H. Thomas James, Vice-Chairman, with a Review of the Report by a Committee of the National Academy of Education (Washington, D.C.: National Academy of Education, 1987), 11.

23. National Science Board Commission on Precollege Education in Mathematics, Science and Technology, *Educating Americans for the Twenty-First Century* (Washington, D.C.: NSBCEM, 1983); Thomas A. Romberg, "The NCTM Standards as a Model for Discipline-based School Reform" in *Telecommunications as a Tool for Educational Reform: Implementing the NCTM Mathematics Standards* (Wye Center, Md.: Aspen Institute, 1992), 13.

24. Romberg, "NCTM Standards as a Model," 13; National Council of Teachers of Mathematics, *Curriculum and Evaluation Standards for School Mathematics* (Reston, Va.: NCTM, 1989), v.

25. National Council of Teachers of Mathematics, *Curriculum and Evaluation Standards*, 2; Romberg, "NCTM Standards as a Model," 16.