

Curriculum and Competence

E. D. Hirsch Jr.

DEFINING ACADEMIC COMPETENCE

In a democracy, a minimal aim of schooling is to bring every child who is not mentally retarded to competence. Recently, in the new economy, even dictatorships have attempted for economic reasons to foster universal education. But long before the information age, indeed ever since the American and French Revolutions, the aim of universal competence—universal literacy and numeracy among the people—has been an avowed educational goal of democracies, where the people rule. Such competence means at a minimum the ability to understand written and oral materials addressed to a general audience; the ability to communicate in speech and writing; and the ability to do math up to prealgebra.

This chapter makes the straightforward claim that it is impossible to bring all children to competence without conveying needed knowledge to them through a coherent, cumulative curriculum in the early grades. Literacy, for example, depends at a minimum on having a wide enough vocabulary to understand newspapers, talk shows, teachers, and vocational training manuals. Having this broad vocabulary depends on having the broad knowledge represented by the words, and broad knowledge cannot be conveyed to

everyone through schooling without schools' offering an effective, cumulative curriculum free of huge gaps and boring repetitions. The failure of our public schools to deliver such a curriculum is the chief reason for the U. S. failure to bring all children to competence.

Several of the distinguished authors of this book favor structural changes to improve school performance, especially changes that impose upon our ill-performing monopoly the element of competition and other free market principles. I also favor the principle of competition in schooling. But efficient markets depend on the free flow of accurate information. We know enough about what works and what does not work to intervene in the educational market in order to prevent malpractice and misinformation from having a distorting influence on the competitive environment—just as false and deceptive advertising does. That is why three of the authors, Ravitch, Walberg, and I, have focused as much on content and pedagogy as on market-style structural change—which is to say we have focused on some of the historical, cultural, and scientific foundations of effective versus ineffective schooling. My task in this chapter is to discuss why it is the case that each and every school in order to be effective needs a specific, coherent, and well-conceived curriculum, plus sound tests that indicate how well that curriculum has been taught and learned.

The immediate policy implications of this thesis lie in its relevance to the state standards movements. Forty-nine of the fifty states have imposed curricular standards exhibiting varying degrees of prudence and effectiveness. It has been argued that some of the existing standards and some of the tests by which they are enforced have actually diminished educational effectiveness and are failing in their avowed purpose of bringing all children to competence. To the extent that these defects exist, they should be remedied. But the criticisms should be met, not be allowed to halt the progress of the standards movement. The standards movement says in essence that every school should deliver a coherent, well-conceived curriculum and should give sound tests that indicate how well the curriculum has been taught and learned.

My chief complaint about the state standards movement is its

failure to achieve this fundamental aim. The so-called content standards are often vague to the point of nullity. Moreover, there is a needless lack of fit between the various content standards among the various states and localities. If all the states agree that young children need to learn about ancient Egypt and the river Nile, then, given the constant movement of children from one state to another, it is pointless and harmful that Egypt should be assigned to grade one in one state and to grade three in another.

Needless to say, a national curriculum—so labeled—is not on the policy table. But state and local curricular guidelines are to be found everywhere. Once a state decides that children should learn certain skills and contents, then the sequence in which they are taught should be subject to negotiation with other states that have decided on similar goals. This would lessen the harmful effects of student mobility and make textbooks more effective. Textbooks would then compete for their effectiveness in delivering a curriculum, not in the pointless novelty of a new sequence for the curriculum. A sequence is not a curriculum. To the extent that state guidelines share similar cognitive goals, there should be a convergence of state standards toward an agreed-upon sequencing of those goals. That is the first policy proposal of this chapter. The second policy proposal is that the high-stakes tests that measure whether the standards have been met should be improved.

THE MYTH OF THE EXISTING CURRICULUM

The curricular chaos of the American elementary school is a feature of our public education that few people have been even remotely aware of, and the growth of that awareness has been one of the origins of the standards movement. We know there is no national curriculum, but we assumed, quite reasonably, that agreement had been reached in the district or school regarding what shall be taught to children at each grade level. The stated reason for preserving the principle of local control of education has been that the localities ought to determine what our children shall learn. But the idea that there exists a coherent plan for teaching content within the local district, or even within the individual school, has been a gravely misleading myth.

That the idea is a myth is not a darkly kept secret. Rather, the idea that there is a local curriculum is accepted as truth by experts within the school system. Recently, a district superintendent told me that for twenty years he had mistakenly assumed that each of his schools was determining what would be taught to children at each grade level but was shocked to find that the assumption was entirely false. He discovered that no principal in his district could tell him what minimal content each child in a grade was expected to learn. He was not surprised when I told him I had received a letter from a distraught mother of identical twins in which she complained that her children had been placed in different classes at the same school and were learning totally different things.

Local districts have produced thick documents that call themselves “curriculum guides” but which, for all their thickness, do not answer the simple question “What specific content are children at a grade level required to learn?” This same defect applies to most of the new state standards. If challenged on this point, an administrator at the state or local level might respond that each teacher at a grade level is encouraged to use a certain textbook. But using specific textbooks does not assure teaching minimal content any more than a thick pile of guidelines. Consider the following research regarding textbook-use in American schools:

Daunted by the length of most textbooks and knowing that the children’s future teachers will be likely to return to the material, American teachers often omit some topics. Different topics are omitted by different teachers, thereby making it impossible for the children’s later teachers to know what has been covered at earlier grades—they cannot be sure what their students know and do not know.¹

GAPS AND REPETITIONS

It might be wondered how it is possible for states and localities to produce lengthy curricular guides that, for all their bulk, fail to define specific knowledge for specific grade levels. Here are some typical instructions. They pertain to first-grade social studies.

1. H. Stevenson and J. Stigler, *The Learning Gap: Why Our Schools Are Failing and What We Can Learn from Japanese and Chinese Education* (New York: Summit Books, 1992), p. 140.

The child shall be able to identify and explain the significance of national symbols, major holidays, historical figures and events. Identify beliefs and value systems of specific groups. Recognize the effects of science and technology on yesterday's and today's societies.

These words disclose a characteristic reluctance of official guides to impinge on the teacher's prerogatives by stating which national symbols, major holidays, historical figures, and historical events the local curriculum makers have in view. But, in the absence of specifics, is there any reason to believe that different teachers will respond to these directions in similar ways? When children from first-grade classes enter second grade, what shared knowledge can the second-grade teacher take for granted among them? What are the "specific" groups that the students became acquainted with in studying the "beliefs and value systems of specific groups"? The word *specific* in such a context carries an unintended irony. Some states have moved toward greater specificity, a hopeful sign.

A few district curricula are more specific but exhibit grave unevenness in thoughtfulness and coherence. Take, for example, the subject of plants and seeds in the best local science curriculum I could find.

- Grade one: "Describe seeds and grow plants from seeds. State three requirements for seed germination and plant growth."
- Grade two: "Arrange illustrations of plants in various stages of development in order from seed to adult."
- Grade four: "Plant seeds and identify and determine the environmental factors responsible for the success and failure of plant development."
- Grade five: "Identify and plot the growth of the seed parts and infer that the cotyledon is food for the living embryo."

The theory behind this sort of repetition is that of deepening through "spiraling." But it is universally experienced by students as boring repetition, as in the oft-heard complaint from students who have been made to read *Charlotte's Web* three times in six grades.

If repetition and boredom are dangers in the “strand” approach to curriculum, an equal danger is the creation of gaps that open up in the spaces between the strands. Frequent repetitions and gaps are the besetting weaknesses of American curricula, and they are made inevitable when the strand approach is compounded with vagueness. Huge gaps are bound to arise. There was no indication even in this topflight local curriculum as to when children were to be introduced to photosynthesis, or that they were to be made aware of simple tools and how they work, or that they would know how to measure physical things in inches, feet, pounds, kilograms, grams, quarts, pints, or cubic centimeters. It might be assumed that the individual teacher would fill in these gaps. But experience has shown this to be an unwarranted assumption. Major gaps in the local guidelines become major gaps in students’ minds—especially among students for whom the school is the only source of academic knowledge.

STANDARDS FROM PROFESIONAL ORGANIZATIONS

To rectify these problems, national organizations were recently commissioned to supply national standards in the different subjects, with notoriously uneven results. The standards for language and literature were rejected by the Department of Education as being entirely vague. The standards in history were denounced by a vote in the U. S. Senate. The curriculum standards in math have become a subject of fierce debate within the mathematics community. And there are three national science standards at odds with one another. Quite apart from these ongoing debates at the national level, the states and districts that have accepted the national standards still face the task of translating them into curricula that overcome the defects of the vague guidelines they are supposed to replace.

How did we achieve this degree of curricular ineptitude, unique in the developed world? As Diane Ravitch has shown in detail, beginning in the 1930s as part of the advance of progressive education in the public schools and colleges of education, there were curriculum revision movements across the land. Over the past six

decades such vague, gap-ridden “conceptual” curricula were developed as a reaction to earlier content-oriented approaches to forming a curriculum. The new curricula have attempted to get beyond the “rote learning” of “mere facts” and to gain unity and conceptual depth by following broad and deep instructional aims. Even the best local and state guides of this type have fundamental weaknesses.

The first inherent weakness is the arbitrariness of the conceptual schemes and classifications that make up all broad curricular “strands” or “objectives.” Such schemes may appear to be deep and comprehensive, but most are quite arbitrary. The large conceptual objectives in each state and district tend to be different from one another, with each state and district preferring its own. Equally striking is the arbitrariness of the different conceptual schemes that curricular experts recently produced for the American Association for the Advancement of Science, the National Council of Teachers of Science, and the National Academy of Sciences, each with its own conceptual scheme.

There is another inherent shortcoming in an overreliance on abstract large-scale objectives (as opposed to “mere” content) as a means of determining a curriculum. These general objectives do not compel either a definite or a coherent sequence of instruction. That is because the large conceptual scheme and the concrete expressions of the scheme through particular contents have a very tenuous and uncertain relationship to each other. A big scheme is just too general to guide the teacher in the selection of particulars. For instance, one multigrade science objective in our superior local district states: “Understand interactions of matter and energy.” This is operationally equivalent to saying “Understand physics, chemistry, and biology.” The teachers who must decide what to include under such “objectives” are given little practical help.

Adequately detailed guidelines help teachers by discriminating between knowledge that is required and knowledge that is merely desirable. The selection of particular important “facts” reduces the total number of facts that a teacher needs to consider essential. Without specifics, disadvantaged students and their teachers play a Kafkaesque game whose rules are never clearly defined. Soon the

unlucky are consigned to slow tracks and never enter the mainstream of learning and society.

THE PROBLEM OF MOBILITY

A systemic failure to teach all children the knowledge they need in order to understand what the next grade has to offer is the major source of avoidable injustice in our schools. It is impossible for a teacher to reach all children when some of them lack the necessary building blocks of learning. Under these circumstances, the most important single task of an individual school is to ensure that all children within that school gain the prior knowledge they will need at the next grade level. Because our system currently leaves that supremely important task to the vagaries of individual classrooms, the result is a systemically imposed unfairness even for students who remain in the same school. That inherent unfairness is greatly exacerbated for children who must change schools, sometimes in the middle of the year.

Add to these academic handicaps the emotional devastation of not understanding what other children are understanding, and add to avoidable academic problems the unavoidable ones of adjusting to a new group, and it is not hard to understand why newcomers fail to flourish in American schools. Then add to all of these drawbacks the fact that the social group with the greatest percentage of school-changers are low-income families who move for economic reasons, and one understands more fully why disadvantaged children suffer disproportionately from the curricular incoherence of the American educational system.

It is often said that we are a nation of immigrants. We are also a people that continues to migrate within the nation's borders. According to the United States General Accounting Office, about one-fifth of all Americans relocate every year.

The United States has one of the highest mobility rates of all developed countries; annually, about one-fifth of all Americans move. Elementary school children who move frequently face disruption to their lives, including their schooling.²

2. General Accounting Office (GAO), "Elementary School Children: Many Change School Frequently, Harming Their Education," GAO/HEHS-94-45, Feb. 1994, p. 1.

In a typical community, the average rate at which students transfer in and out of schools during the school year is nearly one-third: The average rate for Milwaukee public elementary schools is around 30 percent. And among the parents who move, it is those in the lowest income brackets who move most frequently—much more often than middle- and high-income families. This high mobility among low-income parents guarantees that disadvantaged children who will be most severely affected by the educational handicaps of changing schools are the very ones who move most often. In a typical inner-city school, only about half the students who start in September are still there in May. The myth of the local curriculum can be matched by the myth of the local school—if one means by the term *local school* not just a building and a staff but also the students who attend it during the year.³

Student mobility is rarely mentioned in discussions of school reform. That says more about the self-imposed restrictions on our educational thinking than about the urgencies of our educational problems. Any challenge to the principle of an autonomous local curriculum is considered taboo. Hence all the problems that are exacerbated by that taboo, including the deleterious effects of student mobility, receive far less public attention than they deserve.

The term of art for the percentages of transferred students is “mobility rate.” The average mobility rates for the inner city lie routinely between 45 and 80 percent, with many suburban rates between 25 and 40 percent. Some inner-city schools in New York City and elsewhere have mobility rates of over 100 percent. That is to say, the total number of students moving in and out during the year exceeds the total number of students attending the school. “In some of the nation’s most transient districts where some slots turn over several times, schools have mobility rates of more than 100 percent.” The adverse effects of these moves on educational achievement contribute significantly to the low achievement of our

3. Deborah Cohen, “Moving Images,” *Education Week*, August 3, 1994, pp. 32–39. David Wood, Neal Halfon, and Debra Scarlata, “Impact of Family Relocation on Children’s Growth, Development, School Function, and Behavior,” *Journal of the American Medical Association* 270, no. 1334–8 (September 15, 1993).

system as a whole. The General Accounting Office found that many more migrating third-graders were reading below grade level, as compared those who had not yet changed schools. Given the curricular incoherence of schooling even for those who stay at the same school, the fragmentation and incoherence of the education provided to frequently-moving students approaches the unthinkable.⁴

The deleterious effects of mobility should be placed in a particular historical, cultural, and educational context, rather than conceived of as timeless and inevitable. There is strong evidence that the adverse effects of student mobility are much greater in the United States than in countries that use a core curriculum. In a summary of research, Herbert Walberg, citing the work of Bruce C. Straits, states that “common learning goals, curriculum, and assessment within states (or within an entire nation), moreover, also alleviate the grave learning disabilities faced by children, especially poorly achieving children who move from one district to another with different curricula, assessment, and goals.”⁵

Even those who hold strongly to the principle of local control of curriculum might well concede the need for a voluntary agreement about a common sequence in the curriculum—at least in those areas like math and science and basic facts of history and geography that, unlike sex education, are not or should not be subjects of controversy. Other things equal, the principle of the local curriculum is desirable in a democracy. But things are not now equal or effective or fair. Against the principle of local autonomy must be weighed the principles of educational excellence and social fairness. Democratic principles sometimes conflict with one another; none is absolute.

TESTS: THE ROLE OF THE STATES

Statewide content standards are beginning to spawn high-stakes tests that have evoked furious opposition—not without cause.

4. GAO, “Elementary School Children,” p. 6.

5. H. J. Walberg, “Improving Local Control and Learning,” Preprint 1994. Walberg cites B. C. Straits, “Residence, Migration, and School Progress,” *Sociology of Education*, 1987, 60, 34–43.

Under present circumstances, the backlash against curriculum-based tests has been warranted. The policymakers who have instituted these high-stakes tests have made two strategic mistakes. First, they introduced content standards and tests before providing teachers and students with detailed outlines and teaching materials that define what the content standards really are. They have put in place no adequate system for training teachers in the subject matters identified by the content standards. They have failed to do the hard work of deciding which aspects of the content are the most essential to be included in textbooks, teacher seminars, and tests—a lack of specificity and selectivity that has made at least some of the tests less reasonable and fair than they should be. Despite these flaws, the curriculum-based tests (where the state standards are beginning to be specific enough to generate a curriculum, as in California, Virginia, and Massachusetts) are the most promising educational development in half a century.

How these curriculum-based tests should be phased in as criteria for student promotion and graduation is a practical and political question to be decided in a democracy by the representatives of the people. I want to shed light on a technical issue that can be useful in helping to make such policy decisions better informed—that is, the differences and the connections between competency-based tests and curriculum-based tests.

Competency-based tests sample knowledge from a broad range of domains, which enables the tests to exhibit a reliably high correlation between test scores and real-world competencies. Curriculum-based tests are narrower. They try to determine how well specific content standards in a particular domain for a particular age group have been learned. Whereas competency tests indicate overall achieved ability, curriculum tests indicate whether specific knowledge has been gained. The astute reader will perhaps see where I am going—that a well-devised curriculum, monitored by good curriculum-based tests should, over time, extend the breadth of a student's knowledge and thus raise scores on broad-gauged competency-based tests.

Because an indispensable aim of schooling is to increase student competency, the public has a right to demand that results on the

two kinds of test should in due course show a positive correlation. I shall try to explain why good curriculum-based tests, based on good content standards, are the surest and most democratic means of raising scores on competency-based tests and achieving real-world competencies.

COMPETENCY TESTS

An excellent example of competency-based tests would be standardized reading tests such as the verbal portions of the Stanford 9, the ITBS, CTBS, the Nelson-Denny reading test, and so on. Although these are norm-referenced instruments that rank students against one another in percentiles, they can also be scored to indicate a student's grade level of reading comprehension. A score of 5.2 would mean that the student is reading at the level the average student has reached by the second month of grade five. These grade-level calibrations (which have been criticized on various grounds) could be also translated into absolute scores that can be equated over many decades. All the well-established reading tests are valid, reliable, and highly correlated with one another.⁶ What sorts of questions are asked on a standardized reading test that cause it to indicate so reliably academic achievement and readiness? In the earliest-grade versions, there are of course questions about sounds and letters. Later versions include questions about vocabulary, the meanings of individual sentences, and the implications of passages from literature, the natural sciences, the social sciences, practical affairs, and several other domains. How could

6. The intercorrelation of reading tests with one another forms part of the technical literature accompanying the tests. Different tests published by a large company are often "equated" to the other reading tests or test components sold by the company. Researchers have found strong intercorrelations between reading scores on the Armed Services Vocational Aptitude Battery (ASVAB) and the Armed Forces Qualification Test (AFQT) and the various standardized reading tests such as Gates-Maginitie, Nelson-Denny, and the Stanford Tests of Academic Skills. The intercorrelations determined for reading-related skills range between .99 and .87—at the very limits of the reliability of the tests! See B. K. Waters, J. D. Barnes, P. Foley, S. Steinhaus, and D. C. Brown, *Estimating the Reading Skills of Military Applicants: Development of an ASVAB to RLG Conversion Table* (Alexandria, Va.: Human Resources Research Organization, 1988).

such a test, disconnected from any specific curriculum, so reliably calibrate academic achievement, learning readiness, and even real-world competency? One needs to offer not just the ample evidence that this claim is true but to provide a credible theory that explains the strong correlation between reading and general competency.

So I think it will be useful to state some of the theoretical principles that explain why good competency tests in reading turn out to be powerfully indicative of achieved abilities that go far beyond reading. Such a theory has the additional benefit of explaining the potency of curriculum-based tests and illustrating the bond of necessity between curriculum and competence.

1. Reading has been shown to be a process of mentally rephrasing language, rather than being a separate linguistic process. The interpretation of the written word is a reenactment of the interpretation of the spoken word. Many of the conventions used in written language are used in speaking and listening. This mental reenactment of speech explains why reading ability is correlated with general communicative competence—the ability to understand and make oneself understood in oral as well as written speech.⁷
2. Such general communicative competence is required for effective social intercourse in modern society, and is especially critical in schooling, where it forms the basis for understanding the oral and written communications of other people, including teachers.
3. The level of one's reading ability (as reflected in the vocabulary items and passage types on a reading test) predicts the level of one's ability to learn new things. A person learns

7. R. Conrad and A. J. Hull, "Information, Acoustic Confusion and Memory Span," *British Journal of Psychology* 55 (1964): 429–32; D. L. Hintzman, "Articulatory Coding in Short-Term Memory," *Journal of Verbal Learning and Verbal Behavior* 6 (1967): 312–16; M. Naveh-Benjamin and J. T. Ayres, "Digit Span, Reading Rate, and Linguistic Relativity," *Quarterly Journal of Experimental Psychology* 38, 379–51. A general discussion of the underlying "phonological loop" is to be found in A. Baddeley, *Human Memory: Theory and Practice* (Needham, Mass.: Allyn & Bacon, 1998), pp. 52–70 and *passim*.

new things by associating them with things already known. Scoring high on a reading test requires a broad vocabulary that represents broad knowledge that offers multiple points of association for gaining further knowledge. The more you know, the easier it is to learn still more—a principle well established in cognitive psychology.

This is the critical element of the theory. Breadth of knowledge is the single factor within human control that contributes most to academic achievement and general cognitive competence. Breadth of knowledge is a far greater factor, for instance, than socioeconomic status. The positive correlation between achieved ability and socioeconomic status is .422, whereas the correlation between achieved ability and general information is .811. This little-known and momentous fact means that imparting broad knowledge to all children is the single most effective means of narrowing the competence gap through schooling.⁸

4. A score on a test of reading ability shows the degree to which this broad knowledge is readily deployable. A merely passive vocabulary that cannot be marshaled and used critically for reading comprehension is inert knowledge. Psychologists use terms like *accessibility* and *availability* to describe such actively usable knowledge. Accessibility of knowledge is attested to by a person's ability to bring that knowledge to bear in comprehending and analyzing the diverse passages in the test.⁹

8. For these precise correlations see D. Lubinski and L. G. Humphreys, "Incorporating General Intelligence into Epidemiology and the Social Sciences," *Intelligence* 24, no. 1, 159–201. In cognitive science the knowledge-competence principle has become so foundational that it has branched off into different specialties such as schema theory and expert-novice studies. Experts learn new things faster than novices do because of the high accessibility of multiple points of reference and analogy. See, for instance, J. Larkin et al., "Models of Competence in Solving Physics Problems," *Cognitive Science* 4 (1980): 317–48. General discussions may be found in any textbook on cognitive psychology. See, for instance, A. Baddeley, *Human Memory: Theory and Practice* (Needham, Mass.: Allyn & Bacon, 1998), pp. 125–43.

9. See, for instance, E. Tulving, "The Effects of Presentation and Recall of Material in Free-Recall Learning," *Journal of Verbal Learning and Verbal Behavior* 5, 193–97. Baddeley, *Human Memory*, pp. 193–94.

In sum, theory predicts that a good reading test will indicate students' level of communicative competence, their breadth of knowledge, and their ability actively to apply that knowledge to learning new things. Theory further predicts that these competencies will correlate well with job performance and the capacity to be an active citizen because communicative competence and the ability to learn new things are highly important skills in meeting the duties and responsibilities of the modern world.

These predictions are confirmed by massive evidence.

1. Scores in early reading tests predict scores in later reading tests. The more one reads, the more automated becomes the process, and, through reading itself, the broader becomes one's knowledge and vocabulary, and consequently the more readily one understands ever more difficult matter.¹⁰
2. Scores on reading tests predict grades in school. There is a positive correlation between reading scores and academic achievement.¹¹
3. Scores on reading tests predict job performance. Obviously, reading scores do not predict whether somebody can fix your car's engine. But, according to studies conducted by the armed services, reading scores do predict how readily and well a person will learn to fix your car's engine.¹²

10. Anne E. Cunningham and Keith E. Stanovich, "Early Reading Acquisition and Its Relation to Reading Experience and Ability 10 Years Later," *Developmental Psychology* 33, no. 6 (November 1997): 934–45.

11. Sari Lindblom-Ylänne et al., "Selecting Students for Medical School: What Predicts Success during Basic Science Studies? A Cognitive Approach," *Higher Education* 31, no. 4 (June 1996): 507–27. Boris Blai Jr., "The Nelson-Denny Reading Test and Harcum-Earned Academic Averages," Harcum Junior College, Bryn Mawr, Pa., June 1971. Sirkka Gudan, "The Nelson-Denny Reading Test as a Predictor of Academic Success in Selected Classes in a Specific Community College," Schoolcraft College, Livonia, Michigan, January 1983.

12. B. L. S. Scribner, D. A. Smith, R. H. Baldwin, and R. L. Phillips, "Are Smart Tankers Better? AFQT and Military Productivity," *Armed Forces and Society* 12 (1986): 193–206; D. Horne, "The Impact of Soldier Quality on Army Performance," *Armed Forces and Society* 13 (1987): 443–45; J. C. Fernandez, "Soldier Quality and Job Performance in Team Tasks," *Social Science Quarterly* 73 (1992): 253–65; C. Jencks and M. Phillips, eds., *The Black-White Test Score Gap* (Washington, D.C.: Brookings, 1998), pp. 14–15, 75–76.

4. Scores in reading tests predict income. Given the causal connections between communicative ability, learning ability, and job performance, it is not surprising that superior job skill should be rewarded, on average, with superior pay.¹³

THE NEED FOR CURRICULUM-BASED TESTS

The scores on a reading test or other competency test may sometimes be relatively independent of the quality of schooling. One's reading score is better predicted by one's family environment and the amount of reading one has done than by the school one attends. This is a version of the finding by Coleman that influences outside the school are more determinative of academic achievement than influences inside the school. This is not an inevitable sociological law but rather a persistent feature of current American schooling, and it does not hold with the same force in France or Sweden. The gap-closing educational results in these countries remind us that an important purpose of democratic schooling is to help able people overcome accidents of birth and circumstance. I believe that educational policy in a democracy should aim to create a system of schooling in which scores on reading tests depend much more on school influences than they recently have in the United States.¹⁴

Schools can accomplish this egalitarian purpose by making students better readers, that is, causing them to score higher on competency tests, whether or not they come from educated homes. This goal can be reached only by an effective, cumulative curriculum that gradually builds up the knowledge and vocabulary that is being sampled in a reading test. This seems to me a criterion that should be met by state curriculum standards: Will teaching

13. Jencks and Phillips, *The Black-White Test Score Gap*, pp. 445, 489–94 passim; C. Richard Hofstetter, Thomas G. Sticht, and Carolyn Huie Hofstetter, "Knowledge Literacy, and Power," *Communication Research* 26 (February 1999): 58–80.

14. R. Erikson and J. Jonsson, eds., *Can Education Be Equalized? The Swedish Case in Comparative Perspective*, (Boulder, Colo.: Westview Press, 1996). For translated articles on France and data see www.coreknowledge.org, link to preschool, link to "French Studies."

this content provide children with high communicative competence and the ability to learn new things, no matter what their home disadvantages may be?

This democratic criterion means putting in place the very policies that have created the current backlash—setting forth grade-by-grade knowledge standards and monitoring whether that knowledge is being gained, an aim that has won strong support in low-income districts that recognize the democratic effect of this reform.

John Bishop of Cornell has shown that educational systems that require definite content standards and that use curriculum-based tests to determine whether the curriculum has been learned greatly improve achievement for all students, including those from less-advantaged backgrounds. Additional evidence in support of curriculum-based testing comes from the recent finding that gains in reading are directly proportional to the completeness with which a school implements a coherent, content-rich curriculum. Put starkly, a system of coherent standards, coupled with curriculum-based tests, will in fact cause achievement on noncurriculum-based tests to rise. It will result in higher achievement overall and a narrowing of the academic gap between rich and poor.¹⁵

But this change must be instituted wisely, and the critical policy decisions must not be left to technical test makers. Testing companies are very good at creating instruments that have good “psychometric properties,” that is, which rank-order students in a smooth, normal curve. Curriculum-based tests should not exhibit those statistical properties, at least not at first. The tests should mainly dis-

15. John Bishop, *Do Curriculum-Based External Exit Exam Systems Enhance Student Achievement?* (Philadelphia, Pa.: Consortium for Policy Research in Education, 1998). John H. Bishop, “The Effect of Curriculum-Based External Exit Systems on Student Achievement,” *Journal of Economic Education* 29, no. 2 (spring 1998): 171–82. John H. Bishop, “Impacts of School Organization and Signalling on Incentives to Learn in France, the Netherlands, England, Scotland and the United States, Working Paper 93-21, National Center on the Educational Quality of the Workforce, Philadelphia, Pa., November 9, 1993. John Bishop, “The Power of External Standards,” *American Educator* 19, no. 3 (fall 1995): 10–14, 17–18, 42–43. See the three-year Johns Hopkins study excerpted with graphs at www.coreknowledge.org. Level of curricular implementation predicts level of reading gain over three years at multiple sites.

criminate between the students who have gained essential knowledge and those who haven't, with maybe one further category for students who give an abundance of right answers. The earliest versions of the new tests shouldn't rank-order students beyond those three categories—fail, pass, superior. Later on, in a mature, content-based system, such as those Bishop studied, more refined scores might be appropriate.

To grasp the distinction between fancy test items, which aren't appropriate, and plain ones, which are, consider the following examples:

The Civil War ended in

- a. 1864
- b. 1865
- c. 1866
- d. 1867

The Civil War ended in

- a. 1812
- b. 1830
- c. 1865
- d. 1880

Few will doubt that the first question will do a better job of inducing incorrect answers. By including plenty of hard items, test makers can ensure refined, neat rank orderings among students. But it should not be left up to test makers, or even to ad hoc advisory committees, to decide whether students at a particular grade level should have such exactitude of knowledge. That decision should be made and announced in advance by those officials who create the standards and the supporting materials. Curriculum-based exams best serve their purpose, at least at first, by being straightforward and unpedantic.

These considerations lead me to suggest that state education officials should

1. Recognize that we are in a transition period after half a century of content-meager schooling and that state departments

of education must provide the means for teaching and learning the required content standards before too much weight is placed on them. Low stakes before high stakes. Given the historical context, that's only fair.

2. Make public in a very clear and detailed fashion the important aspects of the content standards that are to be emphasized in teacher training, textbooks, and curriculum-based tests.
3. Use the tests as devices to focus effort on productive and important learning that yields centrally useful knowledge and high competence.
4. Grade the straightforward tests generously on a pass-fail basis (with perhaps a "superior" for answering a very high number of straightforward questions) during the transition period while teachers are being trained and appropriate textbooks are being created.
5. Offer, apart from the official, secure tests, informal, no-stakes, year-by-year diagnostic tests that will enable schools to detect knowledge deficits and monitor student progress.
6. Resist any call for a complete test moratorium and give no ground on the basic principle of curriculum-based tests, which are in theory and, as Bishop has shown, also in fact the best route to improved quality and equity.
7. Keep at least a few competency tests in reading, writing, and math. They should carry high stakes (but not unreasonably exalted cut-off points) so long as society agrees that our citizens need these competencies. Well-verified competency-based tests are like those little birds who tell us whether the air in the mine is safe. They reflect the reality principle in education by showing whether competence is truly being achieved.

In short, those states brave enough to have started down this path should continue and improve the policy of using curriculum-based tests, with the stakes gradually getting higher. This is the only known way of achieving the democratic ideal of making the school as effective educationally as the home. That is the appropriate

norm by which content standards and tests should be measured in a democracy.

Those state tests, on the other hand, that are based on no specific content standards mainly increase anxiety without increasing learning. They are no better than commercially available competency tests; in fact they are generally less fair and accurate. For profound theoretical reasons, these skills-tests cannot help schools narrow the achievement gap between groups.¹⁶

In states where good curriculum-based tests are built upon good, specific-content standards, the following can be predicted for current kindergartners and first-graders. By grade seven or eight, when the content-based curriculum has “diffused knowledge” (to use Jefferson’s phrase) and has done much of its compensatory work, academic achievement will have risen for all groups. Higher scores on curriculum-based tests will be well correlated with higher scores on competency-based tests, which will show a significant narrowing of the competency gap between groups. At that point, we shall have moved closer to the ideal of a truly democratic system of education.

16. Besides encouraging time-wasting skills practice on narrow themes, state skills tests offer no theoretical improvement whatever over ordinary competency-based tests, which is what they essentially are. Although they do encourage everyone to work harder in a narrow range, they waste time on empty exercises that cause small gains in general competence and less in equity. They preserve the test-score gap between groups instead of narrowing it because the biggest factor in the competency gap is a gap in general information, which can be narrowed only by a long-range, coherent focus on content. This is another illustration of the importance of basing policy on strong evidence and sound theory.