Setting the Standard: Alternative Policies for Student Promotion

DAVID F. LABAREE
Georgetown University

The widely publicized declines in standardized test scores in recent years have deepened concern about the promotional policies of schools. In this article, David Labaree examines both the theoretical bases and historical evolution of two alternative policies—merit and social promotion. Finding little empirical evidence linking either policy to student achievement, Labaree concludes the article with a series of suggestions for school districts planning to implement stronger promotional standards.

In the last few years, standards for student promotion have become a major topic of discussion among those concerned about public education. Educators, parents, and citizens in general have become worried about the large number of students who are not mastering grade-level basic skills. Increasingly, critics are arguing that one way to solve this problem is to demand that students demonstrate a minimum level of competence before being promoted to the next grade. Such a system of merit promotion is seen as fostering achievement among students, while the social promotion system it is designed to replace is seen as discouraging achievement. Under social promotion policies, students are advanced in response to their social needs—particularly the need to remain with their own age group—rather than in response to their proven ability.

United States public education has been characterized over the years by slow swings of the pendulum between these two alternatives. When common school systems were established in the United States, merit promotion was the rule, but during the course of the twentieth century social promotion came to gain almost universal acceptance; in the late 1970s, the latter policy fell under heavy attack. Recently, for example, school systems in New York City, Chicago, Baltimore, Milwaukee, Richmond, and the District of Columbia have adopted more stringent standards for student advancement. Given the growing concern about student achievement—a concern which has been intensified by the gloomy pronouncements of a number of high-level educational panels—it is likely that other cities will soon fall in line.

For a school board under pressure to do something about poor achievement levels, a promotional standards policy may prove to be irresistible. Such a program seems to of-

Harvard Educational Review Vol. 54 No. 1 February 1984

0017-8055/83/0200-0067$01.25/0

fer a chance not only to improve student performance but also to defuse public criticism by initiating procedures that can be implemented quickly and at minimal cost. Unfortunately, the result is that a number of school boards may find themselves hastening to adopt some form of merit promotion without giving this decision the careful consideration it requires. Before acting, a board should closely examine the consequences of such a program. Is a policy of raising promotional standards really going to bring about the intended increase in student achievement? Are the indirect effects of this policy—on pedagogy, curriculum, organization, finances, politics, and labor relations—desirable or even acceptable?

The aim of this paper is to provide tentative answers to some of these questions within the context of a general discussion of the characteristics and consequences of alternative student promotional policies. It begins with a sketch of the theoretical bases of both merit and social promotion, and includes a brief history of these policies using the Philadelphia school system as a case study. It then examines the empirical literature on the effects of promotional policies on achievement, with special reference to New York City's Promotional Gates Program. Finally, drawing on these observations about the theoretical and empirical implications of alternative promotional policies, the paper concludes with a series of suggestions regarding the implementation of a policy of stronger promotional standards.

The Character and Course of Promotional Policy

Appropriate standards for student promotion have been a concern since the founding of this country's common school system in the early nineteenth century. Prior to that time public education was a small-scale individualized process under which each student advanced through a series of texts at his or her own pace, as determined by recitations with the teacher. In the absence of a comparison group, students experienced neither promotion nor retention but rather a solitary form of forward movement.[1] With the arrival of universal public education in the second quarter of the nineteenth century, student promotions suddenly became an important social issue—the result of the graded structure imposed on the new common school systems at the time of their founding.

Grading was a response to two forms of pressure exerted on the new school systems, one organizational and the other cultural. Organizationally, the sharp rise in the number of students put the common schools under intense pressure to develop a system of instruction which was fiscally, socially, and pedagogically efficient.[2] The result was that they abandoned the inefficiency of traditional individualized instruction in favor of the economies of scale embodied in the simultaneous instruction of an entire class.

[1] Ellwood P. Cubberly, *Readings in Public Education in the United States* (1934; rpt. Westport, Conn.: Greenwood Press, 1970), pp. 73-76; Emile Durkeim, *The Evolution of Educational Thought* (London: Routledge & Kegan Paul, 1977), pp. 252-264. The Lancasterian system—with its combination of group instruction, intense competition, and individualized advancement—acted as a transition between ungraded individualized instruction and graded group instruction. See "Introduction" in Carl F. Kaestle, ed., *Joseph Lancaster and the Monitorial School Movement* (New York: Teachers College Press, 1973), pp. 1-49.

[2] For example, in Philadelphia, public school enrollment jumped from 7,000 to 17,000 in the first year after the founding of common schools in 1836, reaching 45,000 in 1850. See John Trevor Custis, *The Public Schools of Philadelphia* (Philadelphia: Burk & McFetridge, 1897), pp. 18, 22.

Since, under this new technology, the whole class learned the same material at the same time, the students could then proceed on to more difficult material as a group. Individual craft production gave way to large-scale batch production, which in turn led to batch promotion—cohorts of students of similar age and, presumably, similar ability, moving through a progression of educational stages.

Culturally, the new schools were under pressure both to embody and to transmit meritocratic values—particularly the belief that in U.S. society rewards are allocated according to individual ability and effort, and that they are earned, not given.[3] To the extent that a student's rise to each higher stage came as the result of personal achievement, the school system was a hierarchy of merit. Thus, concerns about both efficiency and merit led to the grading of students. The resulting tension centered on promotion. The question was whether the primary unit of promotion was the age-cohort or the individual. The ideal case for educational efficiency has always been to move entire classes through the grades levels like an assembly line with no rejects. The meritocratic ideal has been to promote only those who have reached an acceptable level of achievement.

These alternatives embody different conceptions of the learning capabilities of children and of the goals of public education. Promotion by class rather than by individual student implies that, with relatively few exceptions, all children are capable of learning the same material, although not always at the same time. Schools are seen as attempting to move the great majority of students through their curricula in unison. On the other hand, individual promotion implies that students have widely varied capacities for learning, either because of differences in innate ability or differences in motivation. Schools are seen as trying to select the most able and willing students in order to propel them into higher forms of education while teaching the less capable students at less advanced levels.

This conflict between organizational efficiency and meritocratic values, between the goal of group learning and the goal of individual selection, has been a source of controversy from the time of the first graded schools to the present day. Over the years, three different, though overlapping, core strategies have been adopted in an effort to resolve the problem.

Social promotion. This strategy represents the triumph of efficiency and group learning over merit and individual selection. In its pure form, social promotion means the automatic advancement of all members of an age-cohort from one grade to the next without regard to individual achievement. In the long run it is assumed that achievement levels will converge.

Tracking. In its pure form, this compromise strategy differentiates students into broad categories according to ability. Once this is accomplished, students within each group can either be socially promoted or subject to promotional standards that are differentiated by track. This approach introduces considerable organizational complexity since a variety of curricula must be offered to each age group.

Merit promotion. This strategy represents a stronger emphasis on achievement and selection than on efficiency and group learning. In tracking, the curriculum adapts to the abilities of the students. In merit promotion, the student adapts to the curriculum.

[3] David Tyack and Elisabeth Hansot, *Managers of Virtue* (New York: Basic Books, 1982), pp. 24-28.

Each student is retained or advanced a grade based solely on his or her proven ability as measured against a fixed achievement standard. Organizationally, this strategy leads either to a wide range of ages within each class or to the creation of special classes for those retained, which introduces further organizational complexity and can lead to the development of separate tracks.

Merit Promotion in the Nineteenth Century: Philadelphia

Public school systems in the nineteenth century uniformly adopted merit promotion strategies. For example, the Philadelphia Common School System, founded in 1836, established a meritocratic structure of schooling, and its promotional policies reflected this structure.[4] By 1841 the system already had an exaggerated hierarchical form: while most districts had three levels of schools, Philadelphia had four (primary, secondary, grammar, and high); while most districts had eight elementary grades, Philadelphia had twelve (because of half-year grades for the first four years of schooling). The shape of the school system was that of a pyramid, with a large number of schools at the lower levels, a much smaller number of grammar schools (one for each ward), and only two high schools (one for each sex). Students were selected for admission to each higher level of school on the basis of individual performance on written examinations. This succession of screening procedures culminated in exams for admission to the high schools, though very few students made it that far. Until the very end of the century, public high schools accounted for no more than 2 percent of the students in the Philadelphia school system—primarily because students of all ages chose to enter the workforce or because very few emerged from the rigorous selection process labeled worthy of admission.

This promotional system was geared toward the needs of the city's best students; average students were unlikely to seek admission to the high schools, much less attain it. Yet the system did have, from a certain point of view, its positive aspects. The extreme narrowing at the apex of the educational pyramid meant that retention was too common to be a source of shame and that promotion was perceived as an extraordinary personal achievement. Attaining a high school diploma was so rare that this credential was invested with high status value, and as a result it acted, for some, as a powerful stimulus for achievement. Students were motivated to compete for the honor of attending the high schools, and grammar school principals were motivated to compete for the honor of successfully preparing students for admission.

Nineteenth-century educators felt that this system of meritocratic incentives was in some ways all too effective in spurring student achievement. They worried that it might expose children to mental stress at an early age, thus causing psychological damage. In line with this thinking, the Philadelphia school board in the 1860s launched an all-out attack on the practice of "cramming" for promotional exams. It eliminated some memorization subjects from the high school entrance exam and established maximum time limits for the amount of homework that could be assigned to

[4] The discussion of the history of promotional policy in Philadelphia is drawn from my "The People's College: A Sociological Analysis of the Central High School of Philadelphia, 1838-1939," Diss. University of Pennsylvania 1983.

a student each night—one and a half hours in grammar school, one hour at the secondary level, and none for primary students.

The Rise of Social Promotion

What made the nineteenth-century system of merit promotion work were the limited possibilities for high school education and the resulting ability of the system to motivate the city's best students to compete for admission. By the end of the century, however, these conditions were undergoing rapid change. Aiding in this transformation were two state laws, one in 1887 requiring high schools to accept all qualified applicants, and another in 1895 establishing compulsory attendance for children under the age of thirteen and encouraging the attendance of those between thirteen and sixteen. Thus, in the 1880s, after fifty years with only two high schools, the Philadelphia school board began building new secondary schools, and by 1915 there were thirteen such institutions. At the same time, enrollments at individual high schools expanded rapidly: the student body at the city's oldest high school grew from 500 to 2500 during this period. High school attendance was no longer a rare event.

At a time when most students could not afford the "opportunity cost" of attending high school, selective admissions served the positive function of spurring the ambitions of those who could. But when large numbers of families began to see high school attendance as the natural culmination of their children's education, rigid promotional standards quickly came to be seen as punitive. In 1900 the Philadelphia school board dropped the sixty-two-year-old examination requirement for admission to high schools, and seven years later it abandoned the exam required for promotion in the elementary grades. Students were advanced on the basis of a principal's certification of readiness, which permitted greater flexibility in promotional standards.

As a result of these changes, the district's promotional policy after 1900 made a gradual but steady shift from a merit standard toward social promotion. The clearest indicator of this shift was the upward trend in promotion rates. The rate of promotions in the elementary schools rose steadily from 82 percent in 1908 to a peak of 98 percent in 1945 while the rate for high school promotions rose from 77 to 85 percent during the same period.[5] This relaxation of the promotion standard over the first half of the twentieth century was justified by three related arguments.

First, educators argued that schooling should be structured around the learning needs and abilities of the great bulk of its students rather than the selection and development of the most able. Leonard Ayres, whose book, *Laggards in Our Schools*, led the initial attack on nonpromotion, correctly perceived this restructuring as part of an effort to redefine the basic character of education:

> What is the function of our common schools? If it is to sort out the best of the pupils and prepare them for further education in higher schools, then the most rigorous system, with the severest course of study and the lowest percentage of promotions and the highest percentage of retardation is the best system. But if the function of the common

[5] Philadelphia Board of Public Education, *Annual Reports, Statistical Reports* (Philadelphia: Philadelphia Board of Public Education, 1908-1945).

> school is, as the author believes, to furnish an elementary education to the maximum number of children, then other things being equal that school is best which regularly promotes and finally graduates the largest percentage of its pupils.[6]

In Ayres's view, traditional merit promotional policy measured the performance of the average student against a standard calibrated for the performance of the high-achieving student, with the result that the average student faced a high probability of academic failure. In Philadelphia in 1919 it took students an average of ten years to complete the first eight grades of school.[7] This condition, supporters of the new policies argued, was simply unfair. Moreover, advocates of social promotion asserted that schools should not only adapt themselves to the academic abilities but also to the broader social needs of the average student.[8] In practice this meant a shift from a curriculum-centered school, with its exclusive focus on intellectual development, to a child-centered school, which included concern for the social and emotional development of the student.

Second, educators argued that a zealous policy of nonpromotion seriously impaired the organizational efficiency of the school system. Partly in response to the rapid expansion of schooling at the secondary level, school administrators in the late 1800s and early 1900s were strongly attracted by the possibility of adapting scientific management to help govern their increasingly ungovernable school systems. Cost effectiveness became an important goal, and from this perspective extensive repetition—as reflected in a large pool of over-age students—appeared wasteful indeed. Why should the taxpayers have to pay for ten years of schooling in order to produce an eighth-grade education? Ayres hammered away incessantly at the costliness of retention. He noted, for example, that in 1907-1908 Philadelphia spent almost $900,000 to educate repeaters—more than 20 percent of the total school budget.[9]

Third, educators did not entirely abandon a concern for merit, but they now sought to foster academic achievement not by means of high standards and frequent retentions but by instituting a system of tracking. Interest in tracking developed out of the efforts of educational progressives, who were concerned with preparing students for future occupational roles that were consonant with their different social origins and individual abilities.[10] Differentiated curricula—academic, commercial, manual training—were first introduced into Philadelphia high schools around 1890; later, with the advent of intelligence testing, came special education classes and full-scale ability grouping. Increasingly, merit selection was used as a factor in the process of placing a student within the appropriate track, rather than as a promotional standard.

The Rebirth of Merit Promotion: The Pendulum Swings

In the last two decades there has been a swelling chorus of complaints in this country

[6] Ayres, *Laggards in Our Schools* (New York: Russell Sage Foundation, 1908), p. 199.

[7] Pennsylvania State Department of Public Instruction, *Report of the Survey of the Public Schools of Philadelphia, II* (Philadelphia: Public Education and Child Labor Association, 1922), p. 188.

[8] John Dewey, *The Child and the Curriculum* (Chicago: Univ. of Chicago Press, 1902); Francis W. Parker, *Talks on Pedagogics* (New York: Kellogg, 1894).

[9] Ayres, *Laggards in Our Schools*, pp. 96-97.

[10] Edward L. Thorndike, *Educational Psychology* (New York: Teachers College Press, 1913); William C. Bagley, *The Educative Process* (New York: Macmillan, 1905).

directed toward the practice of social promotion in the public schools.[11] The most frequently voiced criticism is that current promotional policies represent an abandonment by public schools of their once dominant concern with student achievement. The much publicized decline in student scores on standardized achievement tests in recent years has led many people to question whether the schools are doing their job. Why, they are asking, should schools be advancing students to the next grade who have not yet mastered the skills being taught in their current grade? They assert that something is clearly wrong with the structure of schooling when high schools graduate functional illiterates. Social promotion is blamed for much of this deficiency in achievement, for the following reasons:

1. Lowered promotional standards are seen as both reflecting and encouraging the more general decline of standards in American society.[12]
2. Within a school system, a policy of social promotion appears symbolic of a general lack of commitment to student achievement.[13]
3. Setting low minimum achievement levels for promotion is thought to foster low achievement expectations. Critics contend that lowering the "floor" for achievement at a particular grade level leads to a lowering of the "ceiling" as well, while a raised floor leads to a raised ceiling.[14]
4. Promoting students who have not mastered the material for their grade level is perceived as a form of dishonesty. Schools are accused of rewarding students for lack of accomplishment—which instills in them an inflated sense of their own capabilities and teaches them that one can indeed get something for nothing.[15]
5. Rigorous promotional standards are regarded as a positive device for motivating students, parents, and teachers into a sustained effort toward higher levels of achievement.[16]
6. Promoting students according to age rather than demonstrated achievement, opponents of the policy contend, ignores the significant differences in ability and application which mark students within a particular age group. Social promotion sees students as broadly similar in learning capacity and thus seeks to deal with them collectively; but critics charge that students abilities are distributed approximately along a normal curve, which means that schools must make individual discriminations among them.[17]
7. Social promotion is seen as a prime example of a more general problem within the schools, pandering to students. Critics charge that by promoting the unqualified, schools are adjusting their curriculum and instruction to the needs and wishes of

[11] Richard L. Ebel, "The Failure of Schools Without Failure," *Phi Delta Kappan*, **61** (1980), 386-388; Dorothy T. Weathersby, "Are We Failing to Teach Our Children about Failure," *Tennessee Education*, **9** (1979), 3-8; Samuel A. Owen and Deborah L. Ranick, "The Greensville Program: A Commonsense Approach to Basics," *Phi Delta Kappan*, **58** (1977), 531-539.

[12] National Commission on Excellence in Education, *A Nation at Risk* (Washington, D.C.: Dept. of Education, 1983).

[13] National Commission on Excellence.

[14] Ebel, "Failure of Schools."

[15] Ebel, "Failure of Schools."

[16] "Can the Schools Be Saved?" *Newsweek*, 9 May, 1983, 50-58.

[17] Ebel, "Failure of Schools."

> the students when, in fact, students should be adapting to school standards. Critics understand it as the function of schools to lead students, not follow them. They see other examples of this trend—particularly in the proliferation of electives in place of more rigorous academic courses, and in the relaxation of discipline.[18]

As the movement for more rigid promotional standards has gained momentum over the last two decades, it has tended to shift its energies from the attack on social promotion to the establishment of four related types of educational reforms.

Back to basics. At one level this means cutting back the number of electives and special programs in order to increase the amount of instructional time devoted to the traditional academic subjects. At another level, basic skills are typically defined more narrowly—as literacy and numeracy. Thus back to basics is a response to the perception that schools have failed to take the time to teach higher-order academic skills and that schools have failed to teach effectively even the most elementary subjects such as reading and arithmetic. Both forms of basics tend to be stressed in a school system undergoing a shift toward merit promotion.

Minimum competency testing. School systems recoiling from social promotion tend to lean heavily on testing in their effort to raise achievement. Standardized achievement tests—norm-referenced or criterion-referenced—are typically employed to determine whether or not a student meets the minimum requirements for promotion from one grade to another or for high school graduation.

Retention. Typically, students who fail to establish minimum competency in basic skills at the level set by the promotion standard are retained. School systems vary considerably in the degree to which they rely on standardized tests as the criteria for retention, and they also vary over whether the basic skills measured are core academic subjects (usually only in high school) or literacy and numeracy.

Remediation. Usually accompanying a policy of increased retention is a new and intensified program of remediation aimed at bringing the retained students up to a promotable level.

School systems which have adopted some form of more rigorous promotional standard are rejecting the twentieth-century claim for the importance of efficiency in schooling in favor of the nineteenth-century claim for the primacy of merit. There is, however, a great deal more to the reforms in promotional standards than a return to earlier forms of schooling. Too much about the structure and process of education has changed in the course of this century for such a complete return to be possible. For example, while the old system of merit promotion focused resolutely on the needs and abilities of the superior student, the new promotional standards have focused instead on the least able students, on the under- rather than the over-achievers. The aim is to teach basic skills to these students in order to raise them to a minimal level of competency so that a high school graduate will be at least functionally literate.

Another difference between the two promotional standards is that while both have used testing as the criterion for promotion the character of the testing is quite different. During the nineteenth century the critical high school entrance exams that spurred such competitive fervor determined who would be admitted to the high

[18] National Commission on Excellence, pp. 18–21.

school, not who would be retained in the eighth grade. In the days before compulsory education and the decline in dropout rates, a person who failed the exam simply went to work. Now, however, students are compelled to stay until they are sixteen and most remain through graduation. As a result, testing today serves the function of guarding not the entrance but the exit to each grade level.

These differences reflect the radically altered shape of schooling in the 1980s compared with the 1850s. An educational pyramid still exists today, as it did in earlier times, but it has been extended upward well beyond the reach of city school systems. In the 1980s it is the professional schools such as medicine, law, business, and engineering which offer the same combination of exclusiveness and marketability that the city high school did in the mid-nineteenth century. They have the same stimulating effect on college undergraduates that the high school once had for grammar school students.

But the apex of today's educational pyramid—that critical device for motivating students in an ideal meritocratic system—is too far removed from the average student in high school, much less grade school, to provide a realistic goal. Thus modern merit promotional standards lack the positive incentive toward upward mobility that was provided by the old merit system. The incentive that today's public school students have for passing the promotional test is, by contrast, a negative one. They do not want to be held back.

A Case Study in Merit Promotion: New York City

In the past few years school systems in a number of major cities have adopted some form of merit promotion. The most prominent and best documented case of such a policy change is New York City's Promotional Gates Program. This program provides the clearest evidence we have about how a modern merit system can work, and I will examine it in some detail, analyzing the New York City experience in terms of criteria that could be applied to any such policy.

The rigidity of the promotional standard. Is the standard framed in terms of standardized test scores (an inflexible criterion), grades assigned by the teacher (more flexible), or multiple criteria (most flexible of all)?

The validity of the retention criteria. How closely related are the skills being tested to the skills contained in the curriculum and the instruction received in the classroom?

The balance between retention and remediation. Is emphasis placed on holding back low achievers or on providing them with special remedial instruction?

The handling of multiple holdovers. Are there policies defining the number of times a student can be retained and further explaining how to deal with a student who reaches the permissible limit?

The degree of centralization embodied in the policy. To what extent is the central administration strengthened by the process of reforming the promotional system?

The impact of the new policy on student achievement. Do achievement levels rise in the wake of the policy and, if so, should the rise be attributed to the policy itself or to other factors?

The history of the establishment of a system of merit promotions in New York City is unique in several ways. Of all the cities I examined, New York established the most inflexible and test-bound standard for promotion and displayed the strongest commit-

ment to remedial instruction as a balance to retention.[19] In addition, this was the only system which made a determined effort to evaluate the effects of the program. Data are available only for its first full year of operation, so I will be focusing on the period from spring 1981 to spring 1982. The discussion is based primarily on four reports issued by the New York City Office of Educational Evaluation.[20]

The essence of the program is to erect "promotional gates" at the end of the fourth and seventh grades and require students to pass certain tests at these grades in order to move on. The measuring device used was the California Achievement Test (CAT), and the focal skill area was reading. The standard was fixed in terms of grade equivalents: in April 1981 the passing score was 3.7 for fourth graders (one year below the national norm) and 6.2 for seventh graders (one-and-a-half years below the norm).[21] Of the students who took the test in April 1981, approximately 22 percent failed to meet the minimum standard—17 percent of the fourth graders and 26 percent of the seventh graders. With a few exceptions all of these were designated as "eligible" for the Gates program; that is, they were slated for retention. The CATs were administered to these students three times during the following academic year—August 1981 and January and April of 1982. At any one of these times a student earning a score above the minimum could win promotion to the next grade.[22] Overall, 25 percent of the Gates students were promoted in August, 10 percent in January, and 35 percent in the following April, leaving 30 percent to be held over for a second year. More seventh graders became double holdovers than fourth graders—37 percent to 23 percent.

New York City's promotional standards during the first year were extraordinarily rigid when compared with those of other cities. Students who scored below criteria levels on the CATs had to be retained, no matter what their grades were. Fewer than 500 students out of the 24,000 who failed to meet the standard on the April 1981 test were exempted from participation in the Gates program by the Office of Promotional Policy.[23] A single grade-equivalent score for a single skill from the single administration of a single test appears to be a tenuous basis for compelling a student to repeat a year of school. Any achievement test score should be viewed statistically as a rough estimate of a student's true ability and thus is best expressed as a confidence interval rather than as a single figure. By using a cutoff point rather than a cutoff range, New York City guarantees that a number of the students who pass have true scores below the cutoff

[19] In this study I requested information on promotion policy and effectiveness from a number of large northeastern cities whose school systems had recently moved toward a merit promotion policy and whose educational climate was roughly similar to Philadelphia's. The responding cities that met these criteria included New York, Baltimore, Washington, D.C., Milwaukee, and Chicago.

[20] New York City Public Schools, Office of Educational Evaluation, *The Promotional Gates Program: An Analysis of Summer School Participation and August 1981 Test Scores* (New York: New York City Public Schools, 1981); *Promotional Gates Program: An Assessment of Staff Training in the Exemplary Programs, August 1981* (New York: New York City Public Schools, 1982); *The Promotional Gates Program: Mid-Year Assessment and Analysis of January 1982 Test Results* (New York: New York City Public Schools, 1982); *A Final Evaluation of the 1981–82 Promotional Gates Program* (New York: New York City Public Schools, 1982).

[21] The initial regulations mention a plan for replacing the CATs with an in-house, criterion-referenced test, but this proposal has not yet taken effect.

[22] The standard was raised following the January 1982 test date in order to discourage midyear promotions, but at the other two testings the original standard was kept.

[23] The criteria for exemption were relaxed somewhat in the second year of the program to include factors such as other tests and teacher's judgment, which led to a sharp increase in the number of students exempted.

while a number of those who fail have true scores above the cutoff. The August retesting gave students who failed to meet the standard a second chance to pass before being held over, but the point remains that the standard itself is not a valid basis for a pass/fail decision.

Not only is New York's CAT criterion statistically invalid as a basis for promotion decisions, but the instructional and curricular validity of the standard are also questionable. For example, how closely related are the specific skills tested by the CATs to the skills that students have been working on in their individual classrooms? It is hardly valid or fair to evaluate what a student has learned on the basis of a test measuring what he or she has not been taught, or at least has not been exposed to in that particular form. National standardized tests are so abstract in their relation to particular curricula and instructional practices that their validity as measures of student learning should always be suspect.[24] When validity is attained, it is often by means of teaching to the test. Tests designed by a school system to cover the curriculum of that system provide more valid measures, while tests designed by a student's teacher are the most instructionally valid of all.[25] Of course, the latter form of testing fails to provide the uniformity of promotional standards that is generally sought by school districts looking to raise standards. This makes the city-designed, curriculum-based achievement test the optimum compromise between the demands for instructional validity and uniformity of standards.[26]

By far the most positive characteristic about the Gates program was the strong commitment by the school system to provide special instructional support to the students who were retained in grade. Gates students were put into small remedial classes where they received concentrated instruction in carefully selected language and math curricula. The system expended an extraordinary amount of time, effort, and money on the instructional component, underscoring the seriousness of the often-repeated assertion that this program is intended to raise achievement levels not punish underachievers. There was a careful process of curriculum selection, teacher training, oversight, and evaluation; even a citywide summer school was established. The evaluation reports dwell at length on all of these processes, stressing their importance within the overall program.

A chronic problem in a retention policy is what to do with students who have been retained several times. The most challenging case for the Gates program is with seventh graders, where double holdovers are numerous and where students are approaching dropout age. For those who repeat a second time, the year is spent in a Gates Extension Program in which instruction shifts toward the vocational. In the most recent version of the program, students who fail a third time to score 6.2 are simply socially promoted to a high school.

Another consequence of a promotional standards program is organizational centralization. While New York has a turbulent recent history of struggle over community

[24] Robert I. Linn, George F. Madaus, and Joseph J. Pedulla, "Minimum Competency Testing: Cautions on the State of the Art," *American Journal of Education*, **90** (1982), 1-35.

[25] Walt Haney and George F. Madaus, "Making Sense of the Competency Testing Movement," *Harvard Educational Review*, **48** (1978), 462-484.

[26] The related problem of examining only one skill area, reading, is part of the broader problem arising from a basic skills orientation; this will be discussed in the conclusion.

control of the schools, the Gates program has the effect of strengthening the influence of the central administration. The Gates project was a central administration program from the start. In contrast to the decentralized character of many other city programs, Gates was initiated, funded, supervised, and evaluated from 110 Livingston Street. But perhaps the most important centralizing influence comes from the mere existence of a single citywide promotional standard, which forces individual teachers, principals, and community superintendents to fall in line by adjusting instruction to the demands of this standard. In spite of this centralizing influence, the program has engendered surprisingly little opposition from groups supporting community control.[27]

The Evidence Concerning Promotional Policy and Student Achievement

The movement for higher promotional standards received its initial impetus and continues to gain strength from the desire to raise student achievement levels. Proponents argue that competency-based promotion will spur achievement while automatic promotion will have the reverse effect. Since social promotion was slow in establishing its dominance and retention was never completely eliminated, there has been ample opportunity for social scientists to determine which forms of promotion engender the highest level of achievement.

Gregg Jackson's thorough review of the literature in 1975 turned up forty-four studies on the relative benefits of retention and promotion,[28] and an ERIC search in March 1983 unearthed another ten studies completed more recently. Unfortunately, despite the volume of research on the subject, there are no reliable and definitive findings which could serve as the basis for policy. Jackson's conclusion about the literature still holds: "The accumulated research evidence is so poor that valid inferences cannot be drawn concerning the relative benefits of these two options."[29] The problem was not that the studies failed to come up with findings favoring one alternative or the other but that more often than not these findings were invalidated by flawed methodology.

Given the inconclusive character of the evidence, what, if any, contribution can the empirical literature make to the current debate about promotional standards? If we consider the stands taken by the writers of the six major literature reviews published in the last ten years we find that not one of these writers adopts a position in support of retention. Three remain neutral on the policy question,[30] while one, in a report prepared for the Philadelphia school system, mildly favors social promotion,[31] and the remaining two strongly support social promotion.[32]

[27] At least such opposition has not been reported in the *New York Times*.

[28] Jackson, "The Research Evidence on the Effects of Grade Retention," *Review of Educational Research*, 45 (1975), 613-635.

[29] Jackson, "Research Evidence," p. 627.

[30] Jackson, "Research Evidence"; Steven Selden, "Promotion Policy," in *Encyclopedia of Educational Research*, ed. Harold E. Mitzel, III (New York: Free Press, 1982), pp. 1467-1474; "The Literature on Social Promotion Versus Retention," Unpublished Paper, Southwest Educational Development Laboratory, Sept. 1981.

[31] Robert G. Reiter, *The Promotion/Retention Dilemma: What the Research Tells Us*, Report No. 7416 (Philadelphia: Office of Research and Evaluation, School District of Philadelphia, 1973).

[32] Sidney Thompson, *Grade Retention and Promotion* (Burlingame, Calif.: Association of California School Administrators, 1980); Wadi D. Haddad, *Educational and Economic Effects of Promotion and Repetition Practices*, World Bank Staff Working Paper No. 319 (Washington, D.C.: World Bank, 1979).

Evidence from the New York Gates Program

School systems which have instituted a sudden toughening of promotional policy in recent years have established the conditions for a series of natural experiments by which the effectiveness of such policies in raising student achievement can be tested. To my knowledge, however, only New York City has taken full advantage of this opportunity to conduct suitably rigorous evaluations.

The school system produced an evaluation of the results from each of the three tests administered during the first year of the program,[33] but these studies have not established that the policy had a significant impact on achievement. The first two reports are inconclusive because of serious methodological deficiencies, and the more rigorous final report shows no net gain in achievement that is attributable to the new promotional policy.

Any attempt to reach reliable conclusions about the effects of the Gates program on student achievement must first rule out as invalid three alternative explanations for any observed rise in such achievement. Two of these—maturation and prior achievement level—were identified by Jackson as factors which must be taken into consideration in any study of the impact of promotional standards; the third—regression—arises whenever a marginal subpopulation is retested.

Maturation refers to the expectation that students in school will on the average increase their level of achievement over time whether or not they are involved in a special program. The question, therefore, is not whether students in the Gates program made gains but whether their gains were significantly greater than those made by socially promoted students over the same period of time. To answer this question the evaluators must establish a control group of non-Gates students for the purpose of comparison. In addition, if the comparison of final achievement scores between the Gates students and the control group is to be valid, one must adjust these scores to take into account differences in prior level of achievement. Students with higher pretest scores are likely to have higher posttest scores as well, regardless of their participation in the Gates program. A statistical adjustment of the scores permits a comparison of the net gain in achievement due to each promotional policy.

A third source of invalidity in evaluating the effectiveness of the Gates program is regression to the mean, which occurs because of the statistical properties of the testing procedure. Since a CAT score is merely a point estimate of a student's true achievement level, the score will fluctuate from one test administration to another within a predictable probability range. Thus if the lowest-scoring group of students are tested, their scores on average will regress toward the mean of the entire group, which in this case means they will rise. This would occur even if their true achievement levels were unchanged because, in effect, there is nowhere for the fluctuating scores to go but up. The average Gates student gained about five months between the April and August test dates, but only part of this gain is attributable to instruction; the rest is due to regression.[34] Put another way, 25 percent of the April holdovers passed the August test; but the effect of the Gates summer school on this figure is unknown since some of these

[33] New York City Public Schools, *Analysis of Summer School Participation; Mid-Year Assessment and Analysis; Final Evaluation*.

[34] New York City Public Schools, *Analysis of Summer School Participation*, p. 10.

students would have passed even if the retest had been given a few days after its original administration. It is possible to adjust test scores for regression and the final Gates report does so, but (as the report notes) the validity of these adjustments is also open to question—especially in a population subject to periodic attrition such as the Gates group. Once again a control group provides the most secure way of eliminating this explanation of achievement; since retest scores in both groups would be affected by regression, one could attribute the difference between them in net achievement gain to differences in policy.

If a control group can help clarify the real gains in achievement, then the question becomes how to construct the control group. In the ideal social experiment, students are randomly assigned to the experimental or control groups. However, it would be highly unethical to arbitrarily assign some students to be retained in grade. In the absence of pure experimental conditions, the evaluators constructed a comparison group from historical data. This group consisted of those students in grades four and seven from the year prior to the initiation of the program who scored below the Gates minimums on the CATs that year. Under the old promotional policy, 22 percent of these students were retained while the remainder were promoted to grades five and eight. A comparison of the Gates and control students allowed for a reasonably good test of the effects of retention versus social promotion.

Unfortunately, comparison group test scores are available only for April 1980 and April 1981, since before Gates the CATs were given just once a year. This means that the evaluations could provide no comparative data for the August 1981 and January 1982 test results. Both of these reports show sizable gains in student achievement, but without comparative information there is no valid basis for attributing these gains to the Gates program—they could just as easily be the result of extraneous causes such as maturation, regression, or prior achievement. Neither the August nor the January report makes strong claims for the data presented, and the latter document even warns about some of the problems in interpreting the results.

A more rigorous analysis was made of the April 1982 test results. The overall gains in achievement registered by Gates students after the first year were heartening. Even after adjusting for regression—but not for other alternative explanations—fourth graders who qualified for the program in April 1981 gained an average of seven months by April 1982, rising from 3.4 to 4.1; seventh graders gained a full year, rising from 5.4 to 6.4.[35] When a comparison group is introduced—thus controlling for both regression and maturation—the picture becomes more complex. Students who spent a full year in the program—and thus were still in grades four and seven in April 1982—were matched with students from the comparison groups who likewise had been compelled to repeat those grades. Gates students who were promoted in August or January into grades five and eight were matched with students from the comparison group who were socially promoted to the same grades.

The comparison group as constituted therefore controls for both maturation and regression. But in order to produce a valid test of the impact of the Gates program on student achievement, the initial test score of each student must also be held constant. Analysis of covariance is a technique which can accomplish this task. Table 1 shows

[35] New York City Public Schools, *Final Report*, Table 27.

TABLE 1
Reading Achievement by Gates and Comparison Group Students

Grade	Group	Pretest Date	Posttest Date	N	Observed Mean Posttest Scale Score	(S.D.)	Adjusted Mean Posttest Scale Score[a]	Grade Equivalent
Four/five	Gates	April 1981	April 1982	6,924	422.7	(33.0)	423.3	4.1
	Comparison	April 1980	April 1981	6,914	420.6	(33.1)	420.0	4.1
Seven/eight	Gates	April 1981	April 1982	8,659	491.8	(40.3)	491.5	6.4
	Comparison	April 1980	April 1981	10,214	494.6	(39.8)	494.2	6.5

[a]Within-grade analyses of covariance were performed to adjust posttest scores; these scores were adjusted to account for some of the differences in pretest levels.

Source: New York City Public Schools, Office of Educational Evaluation, *A Final Evaluation of the 1981-1982 Promotional Gates Program* (New York: New York City Public Schools, 1982), Table 28.

the comparison between the posttest CAT reading scores of the Gates and control groups when they are adjusted for pretest scores. This procedure statistically approximates an unbiased experimental design for testing the effectiveness of promotion vs. retention.

Unfortunately the adjusted posttest scores for the Gates group and the comparison group are virtually identical. The bottom line is that the students retained under the Gates program appear to have experienced no net gains in CAT scores in excess of the gains experienced by the low-achieving students who were retained or socially promoted under the old system.

Thus one is forced to conclude that there is no evidence that retention and remedial instruction under the Gates program produced any gains in achievement which had not already been produced in the absence of these interventions. Considering how much effort was expended under this program to boost achievement in the Gates group, this finding is quite disheartening.

Of course none of this constitutes proof that the program is ineffective; all that can be said at this point is that its effectiveness remains to be demonstrated. Quite possibly the program will prove more effective over time; it may have a long-run effect on students rather than a short-run effect; it may have an effect on learning that is not measurable by the CATs; and it may have its most significant effect by stimulating the achievement levels of students who surpass the promotional standard rather than those who do not. But such judgments must await better evidence.

Promotional Standards: Proven and Predicted Effects

The national movement toward raising student promotional standards is rooted in a deep concern about achievement. Educators, parents, and the general public are frightened by the widely publicized declines in standardized test scores in recent years and by the growth in the number of high school graduates who have failed to master basic skills. A policy of merit promotion offers a way out of this dilemma by promising to increase the academic demands which schools place on students and to motivate students to meet these demands. Since the decline in achievement is seen as the result of a relaxation of academic standards, it is assumed that an increase in achievement

can be brought about by raising the minimum level of competence required to advance from grade to grade.

The relationship between promotions and performance, however, appears to be more an article of faith than a proven reality. Research evidence on the subject is wholly inconclusive. Out of more than fifty studies of the relative impact of promotion and retention on student achievement, the large majority had a methodological bias which favored one policy or the other. Under these conditions the only significant finding would be one which runs counter to the bias. However, none of the studies produced such a result; instead, results mirrored methodology. The few studies with an unbiased design produced contradictory results. Thus, school systems which raised promotional standards in the last few years did not do so on the basis of this policy's demonstrated effectiveness.

The recent elevation of promotional standards in school systems across the country has created a series of natural experiments in which the impact of the program could be tested. Unfortunately only in the New York Promotional Gates Program (among those programs examined in this study) did evaluators attempt to exploit this situation fully. The final report on the policy's first year in New York showed that most retained students made significant achievement gains during the year; but when the researchers established controls for alternative explanations, these gains vanished. Low-achieving students promoted or retained under the more relaxed standards of the old promotional policy raised their achievement levels in one year by the same amount as the Gates students.

The only conclusion one can draw from the current empirical literature is that there is no valid evidence which demonstrates a significant difference between promotion and retention in their impact on the low-achieving student. Of course the inability to prove a difference in the effectiveness of these policies does not necessarily mean that no such difference exists. Empirical research is conducted according to conservative rules which require that treatments be considered ineffective until proven otherwise. Under these conditions it takes a large number of carefully controlled studies before clear trends can emerge.

The accumulated research evidence, then, should give pause to the school administrator who is planning to raise promotional standards, for the assumption which underlies such a move—that promotional policies are related to achievement—has never been empirically verified. Given the inconclusiveness of the empirical data, the administrator is forced to consider other grounds for making a decision about whether to proceed or not. A likely source of help in such a choice is theory. While we do not know in practice whether such a merit promotion policy is effective in raising achievement levels, there are some theoretical grounds for thinking that it might be. If a policy of raising promotional standards does indeed raise student achievement, it is likely to be for the following reasons.

Fear of retention. Such a policy may turn out to have a significant effect in motivating a student to achieve, and also in motivating the student's parents and teachers to help promote such achievement. In the nineteenth century, merit promotions encouraged students to look ahead to the chance of reward; the same policy today encourages students to look over their shoulders to the possibility of retention. This negative motivation may well be equal in effectiveness to the positive motivation of an earlier time,

but it will most likely influence a different type of student. In the 1980s it is the low-achieving students who are likely to respond to the stimulus since they are the population at risk of retention. In particular, the students most likely to be spurred into action by a merit promotion policy are those receiving a midyear letter announcing that retention will occur unless performance improves. One can imagine such a letter galvanizing parents and teachers as well, with potentially beneficial results for the student's achievement.

Several implications of this motivational system should make an administrator cautious. First, retention is only effective as a motivating device for students to the extent that they find it distasteful. Reasons for such distaste include the unhappiness at being separated from classmates and the shame at being labeled slow. If students feel this way in anticipation of retention, is it not possible that being compelled to experience retention might have harmful effects on their personal adjustment? Of course, proponents of retention policies argue that the policy is not in fact punitive but remedial. The Gates program literature reinforces this notion by referring to the process of failing to meet the promotion standard as "becoming eligible for the Gates program." Yet one cannot have it both ways. If retention is a strong motivating device, then retentions are likely to be fewer, but the students retained are more likely to experience it as punishment. If retention is a weak motivating device, the effect on the student is likely to be more remedial than punitive, but the number retained is likely to be large. No school system wants to make retention unpleasant simply to scare students into passing. The thrust of most of the merit promotion policies studied in this paper has been to make the holdover year a fruitful and pleasant experience. I would argue, however, that such laudable efforts have the effect of undercutting some of the motivational power exerted by retention.

Second, while the fear of retention may motivate the low-achieving student, it is likely to have little or no effect on the average or superior student whose scores are comfortably within the passing range. Therefore, this is not a strategy aimed at raising the minimum level of all students.

Third, the focus on motivation assumes that the problem of underachievement derives from lack of incentive. To the extent that poor test scores are the result of such factors as class background, racial discrimination, family conditions, and test invalidity, the student's motivation is irrelevant, and retention will not spur the student to higher achievement.

Fourth, the news that a child is in danger of failing is likely to have an effect on most parents, but the way in which this effect is transmitted to the child may vary considerably. Parents who interpret the problem as academic may seek to help the student with his or her work, but those parents who interpret the problem as disciplinary may be more likely to punish the student. At home as at school, merit promotion poses a choice between remediation and punishment.

Enhanced remedial instruction. If raised promotional standards do have an effect on achievement, it will be largely the result of the enhanced remediation which, in recent years, has tended to accompany it. Retained students may experience smaller classes, specially trained and motivated teachers, new curricula and more supervisory interest than they were accustomed to in their regular classrooms. School systems have a strong incentive to stress the instructional component of retention in order to under-

score the therapeutic rather than punitive aim of the policy. The intense public and political interest in raising promotional standards may turn out to be a very effective lever for prying loose public funds to pay for this increased level of instruction. In New York the school system succeeded in acquiring a sizable initial commitment of funds from the city for raising standards, most of which went to pay for remediation. Unfortunately, this investment did not appear to pay off in the form of immediate achievement gains.

Focusing attention on achievement. Even if fear of retention and remedial instruction are not effective, raising promotional standards may have a positive impact on student achievement simply as a slogan. Such a slogan could serve as a rallying point for school people interested in emphasizing achievement within the schools by a variety of means, in addition to or even apart from promotional standards. In a report written on promotion and retention for the Philadelphia schools, Robert Reiter sees such a value in a strict retention policy even though his reading of the literature shows social promotion to be superior in practice:

> At this point in our School District's history, it appears that another swing of the promotion-policy pendulum—back toward stricter requirements—might serve as a slogan or symbol under which our zeal for effective education can be renewed. Its slogan value is not destroyed by the fact that a strict retention policy in itself has been found somewhat less effective than a policy favoring social promotion.
>
> Even if research has found it to be less than ideal, no slogan can be "all bad" if its use as a rallying cry indirectly facilitates the really effective classroom conditions under which each child is stimulated to attain his own highest possible level of attainment.[36]

Simulating achievement by the use of tests. It is possible that a policy of raised promotional standards could improve test scores—thus giving the impression of progress—without affecting real achievement. To the extent that a school system devotes time and effort to train students for a particular test, it may raise scores but neglect broader educational objectives. Ideally, schools seek to improve achievement and then measure the improvement with a test. But as soon as promotion becomes contingent on a test score, it may turn out to be more efficient to work on improving the test score and then to attribute increases to gains in overall achievement. Thus, the strongest argument for not relying on a single test as the promotional standard is the wish to keep the tail from wagging the dog.

Suggestions for Raising Promotional Standards

At this moment the tide is moving toward high promotional standards throughout the country. Many school systems have already adopted such a policy, and many who have not probably will do so soon. Under these conditions it may not be realistic to close this paper with a discussion of whether a school system should adopt tougher standards or stay with social promotion. The trend toward the former is so strong that even in systems which have not changed formal promotion policy we often see retention rates rising as a result of informal adjustment. Given this situation, it would be useful to suggest how a policy of raised promotion standards could be implemented, drawing on

[36] Reiter, *Promotion/Retention Dilemma*, pp. 19-20.

the experience of other school systems and reflecting the concerns expressed earlier in this paper.

A flexible promotional standard. At a bare minimum this means not relying on a single score of a single test, as New York does. In the interest of being less punitive and more suited to the needs of the individual student, the standard should be constructed from multiple measures—including curriculum-based tests and teacher evaluations—and should leave room for appeal to higher authorities. Examples of such policies are found in Milwaukee and Chicago.

A valid measure of achievement. Since the process of learning for each student is located within a particular curriculum and a particular mode of instruction, the most valid measure of that student's achievement is the one which best reflects the special character of this learning process. The model for such validity is a teacher's individual evaluation of a student (although this validity is obtained at the expense of citywide uniformity); the least valid measure is the most uniform, a nationally distributed standardized test. Between the two extremes is a city-designed achievement test geared to the curriculum in use. It should be noted that validity is just another word for fairness, and thus concern over this issue should not be limited to methodologists. The moment a standard for promotion is adopted, its fairness will inevitably come under intense scrutiny from parents—and possibly the courts as well.

A rigorous evaluation of program effectiveness. Raised promotional standards are usually put in place under conditions in which much has been promised and much is expected. People inside and outside the system want to see achievement levels go up, and quickly, as a result of the new policy. The temptation is great to give people what they want by presenting only the most favorable data, by failing to employ statistical controls, and even, perhaps, by inflating scores. One way around this problem is for the interested parties to agree in advance on a method of evaluation and on what findings will constitute success or failure. If the program simply does not work, there should be contingency plans for changing or scrapping it.

More than just basics. If students in grade school have difficulty developing a basic competency in reading and math, then they should receive special help in these areas at the expense of other subjects. The same should hold for high school students lacking functional literacy skills. If we take these ideas about correcting learning deficiencies to the logical extreme, however, we will boil the entire curriculum down to its most basic level and, in the process, produce new kinds of deficiencies. One would be a deficiency of interest, since time in school would increasingly be spent on narrowly focused exercises and drills. Another would be a deficiency of breadth and complexity, while ideally schooling should be expansive and challenging.

Include the average student. While concentrating on raising the level of the low-achieving student to a minimum competency, we must not forget the achievement needs of the average student. Minimum competency testing can easily lead to a pass/fail mentality in which those who pass begin to coast, since they feel that no more is expected of them. If higher promotional standards are adopted, it should be as part of a much broader orientation toward high achievement for all students. Without this, a policy of raising standards for the poorest students can have the ironic effect of debasing standards for the rest of the class.

Do not blame the students for the failures of the schools. If we hold students accountable for their own academic performance by retaining those who do not make the grade, we should also hold teachers, principals, and administrators accountable for providing these students with the necessary tools. This kind of accountability is considerably more difficult to implement organizationally than is a promotional standard, but it is a necessary component if we really want to raise achievement rather than just assign blame.

Emphasize instruction over retention. Retention should not be used as a threat but as an opportunity for providing intensified remedial help to those students who most need it. This extra instruction will cost; a full-fledged promotional policy is not a cheap solution to the problem of student achievement. But if taxpayers are convinced that it is part of a serious attack on the problem, they may be willing to provide the necessary support. It is all too easy in the midst of establishing a promotional standard to forget about the special instructional needs created by these standards. Retention puts students on the slow track, and only instruction can get them out of it.

Effective schools. Ultimately, what matters most to student achievement is not one promotional policy or another but the overall effectiveness of the schools in carrying out their mission. Milwaukee's Project RISE (Rising to Individual Scholastic Excellence) is an example of a broad-based program which puts together many of the suggestions made here, and does so in a way that makes promotional policy peripheral rather than central.[37] Beginning with the firm belief that the school by itself can make a difference with the low-income low-achiever, RISE systematically emphasizes all of the factors which its organizers see as characteristic of a truly effective school: grade-level achievement expectations for all students, an orderly learning climate, instructional leadership by the principal, basic skill orientation, frequent inservice training, the establishment of curriculum objectives, regular homework, student identification with the school, heterogeneous ability grouping, direct and structured instruction, concentration of time on task, and a commitment to mastery learning.

Conclusion

This study has focused on merit promotion and social promotion, the two methods used to move students through the graded structure of U.S. schooling. These systems differ in their degree of emphasis on achievement and in their assumptions about student capabilities. Merit promotion, both the old and new versions, is strongly oriented toward spurring achievement, while social promotion tends to place achievement at a lower priority than such concerns as social adjustment and continuous progress. At the same time merit promotion is based on the expectation that students have widely varying degrees of ability, while social promotion perceives students of the same age as having relatively uniform capacities for learning. In this sense the two systems can be seen as mirror images of each other: merit promotion combines elevated expectations about achievement with hierarchical notions of ability, while social promotion combines lower expectations for achievement with egalitarian assumptions about ability.

[37] Milwaukee Public Schools, *Project RISE: Rising to Individual Scholastic Excellence—A Guide to School Effectiveness* (Milwaukee: Milwaukee Public Schools, 1982).

Both systems foster the belief that there is a strong positive association between individual differentiation and excellence, between equality of skills and mediocrity of performance.

Although this belief has dominated U.S. public schooling from its earliest days to the present, alternative models of education do exist which challenge it. Perhaps the most influential such alternative is provided by Benjamin Bloom through his notion of mastery learning.[38] Bloom not only argues that students are broadly similar in their capacity for learning, thus denying the hierarchical assumptions implicit in merit promotion, but he also argues that their capacity extends to complete mastery of the knowledge we want them to acquire, thus denying the minimalist expectations implicit in social promotion. He sees no contradiction between equality and excellence because he attributes the wide variations in student performance to instructional failure—the failure to focus on each student's areas of individual need—rather than to the inability to learn.

In the absence of evidence clearly defining one form of promotional policy to be the most effective, the choice of merit promotion or social promotion—or some alternative program such as mastery learning—must be made on the basis of social values. If we do not know which policy provides a system of instruction that is technically superior, we must at least choose a policy whose implicit values are congruent with our own. Any policy that is implemented, in whichever direction it leans, will involve critical value choices whose consequences will be felt for a long time to come.

Funding for this paper was provided by the Citizens Committee on Public Education in Philadelphia, an independent non-profit group promoting citizen action in support of quality public education. I am grateful to Richard de Lone, Debra Weiner, and Norman Newberg for their comments on an earlier draft. Many of the ideas in this paper first emerged in discussions with Norman Newberg.

[38] Bloom, *Human Characteristics and School Learning* (New York: McGraw-Hill, 1976).