

The Manufactured Crisis



*Myths, Fraud, and the
Attack on America's Public Schools*

David C. Berliner • Bruce J. Biddle

1995



Addison-Wesley Publishing Company

Reading, Massachusetts Menlo Park, California New York
Don Mills, Ontario Wokingham, England Amsterdam Bonn
Sydney Singapore Tokyo Madrid San Juan
Paris Seoul Milan Mexico City Taipei

Chapter Two

Myths About Achievement and Aptitude



This chapter is concerned with a key tenet of the Manufactured Crisis. Critics of American schools have argued that student achievement has declined sharply, that American students now lag seriously behind students in other Western countries, and that these facts are confirmed by massive evidence. America's teachers and schools are failing the nation, they say, and America is in danger of falling into the ash can of history. As we shall show, these assertions are errant nonsense.

Americans are, of course, concerned about the performance of their schools. Such concerns are legitimate and should be encouraged. However, concerns about education and plans for its improvement should be based on an honest and informed evaluation of available evidence. In this chapter we shall use many data sources that lead to clear but perhaps surprising conclusions—that, on average, today's students are at least as well informed as students were in previous generations and that education in America compares favorably with education elsewhere.

What makes these conclusions surprising is that the critics have so often stated otherwise. But repetition of false claims does not make those claims true. Let us therefore restate the major myths about achievement and aptitude that critics have used and look carefully at the evidence that refutes those myths.

MYTH . . . *Student Achievement Has Recently Fallen Across the Nation*

For the first time in the history of our country, the educational skills of one generation will not surpass, will not equal, will not even approach those of their parents.

—Paul Copperman (cited in *A Nation At Risk*, National Commission on Excellence in Education, 1983, p. 11)

Average achievement of high school students on most standardized tests is now lower than 26 years ago when Sputnik was launched.

—*A Nation At Risk* (1983, p. 8)

[From 1950 to 1989] we probably experienced the worst educational decline in our history. Between 1963 and 1980, for example, combined average Scholastic Aptitude Test (SAT) scores—scores which test students' verbal and math abilities—fell 90 points, from 980 to 890.

—William Bennett (*The Devaluing of America*, 1992, p. 55)

Shortly after he became secretary of education in the Reagan administration, William Bennett held a televised news conference. Fortified with impressive visual aids, he claimed that student achievement had declined greatly in the nation. Evidence from objective tests supported this claim, he said. A CRISIS was at hand. Our schools were failing, our youth were becoming ignoramuses, our industries were losing their competitive edge, our country was in *danger*. And given that Secretary Bennett spoke from the "pulpit" of the White House, the press dutifully reported his claims as if they were gospel. Moreover, Secretary Bennett was not alone in making these claims. They were also made in *A Nation At Risk* and other reports that were then being commissioned by government and industry. Reading each other's flawed reports, a host of critics have since chorused that student achievement has declined "massively" and that this decline is confirmed by "many" different standardized test records.

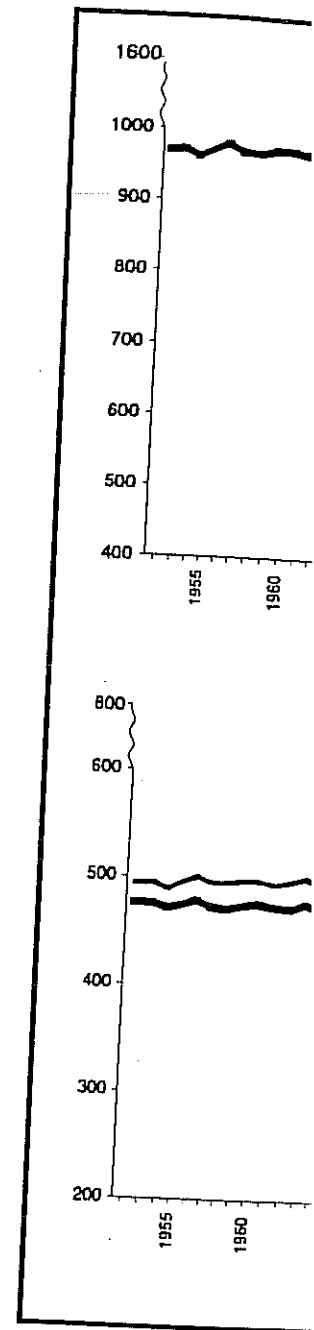
Although we are not sure how appropriate it is to use standardized test data to judge the performance of schools, such tests *do* provide hard, objective evidence that seems relevant to claims about achievement. Standardized tests appear to be rocks of stability in a sea of unanchored opinions, and it is small wonder that people worry when critics make alarming assertions about "massive" declines in student achievement.

We confront these assertions with a novel approach; we review here the actual *evidence* generated by standardized tests. From that evidence, readers will learn that standardized test data reveal *no* recent drop in student achievement; indeed, many of the tests indicate modest recent *gains* in students' knowledge. In fact, we know of only *one* standardized test that ever generated falling aggregate test scores—the Scholastic Aptitude Test (SAT)—but, as we shall show, that decline had nothing to do with average student achievement. In fact, when analyzed correctly, the SAT data also reveal a pattern of achievement *growth*! In short, the real evidence indicates that the myth of achievement decline is not only false—it is a hysterical fraud.

The Scholastic Aptitude Test. We begin with the SAT story. To be sure, *aggregate* total SAT scores obtained by the nation's high school seniors fell between about 1963 and 1975 (see Exhibit 2.1). Moreover, that decline came about because aggregate scores fell for both parts of the SAT—the parts that measure, respectively, verbal and mathematics achievement.

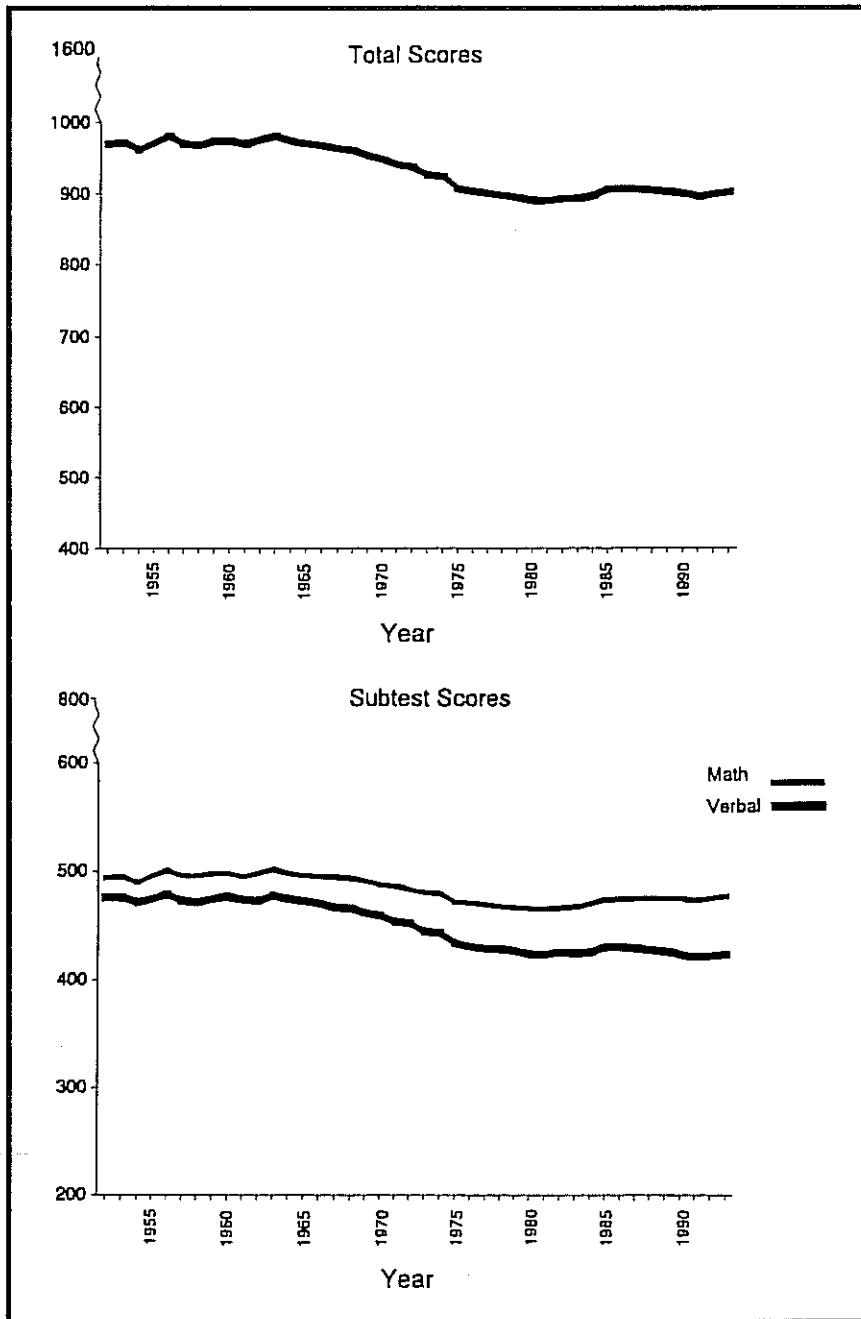
This decline has been cited as evidence that our schools were failing by many people—such as William Bennett—who were either ignorant or

Exhibit 2.1 Average



—Sources: Donlon (1984) and The C

Exhibit 2.1 Average SAT Scores by Year



—Sources: Donlon (1984) and The College Entrance Examination Board (1993).

wanted to rubbish public education. These doomsayers argue that such a "huge" decline in SAT scores—ranging from 60 to 90 points, depending on the years used to compute the figure—was sure proof that the nation was in trouble. And to this day the press continues to report annual figures for aggregate national SAT scores (sometimes citing shifts of one or two points in those scores) as if they were a report card that can tell Americans something about the effectiveness of their schools.

This is all nonsense. Such small shifts in SAT scores are not meaningful, and by themselves aggregate SAT scores provide no information about the performance of American schools. To understand this, one must learn a bit about the SAT and its limitations.

First, how are SAT scores generated, and what does an SAT score truly mean? The SAT is taken by high school seniors and for some years has required those students to answer 138 multiple-choice questions, 78 concerned with verbal materials, 60 focused on mathematics. (It has *not* examined student knowledge of history, the sciences, the arts, the humanities, foreign languages, the social sciences, or other important subjects that high schools also teach.) Although new questions are developed for the SAT each year, those questions have been carefully checked so that each new edition of the SAT has been presumed to be equivalent to the first "standardized" form of the test, which appeared in *April of 1941*. If we believe the critics, then, the quality of the entire twelve or more years of schooling to which American students are exposed could be judged by means of a 138-item test requiring only a few hours to complete, and composed of questions designed to assess only a narrow range of student knowledge against the standards of fifty years ago!

Second, scores for the SAT are not reported as numbers of right answers but are "converted" (through obscure rules) in such a way that the scale scores earned by each student for each subtest range from 200 to 800 points. This means that the total SAT score for an individual can range from 400 to 1600 points, but this huge range of scale points actually represents a much smaller range of correct or incorrect answers. Thus, in the middle of the SAT range, a difference of one correct or incorrect answer will generate about *ten* points of difference in the SAT scale score. You may think this is absurd, but there is actually more confusion. The relation between number of correct SAT answers and scale scores also varies, depending on the number of right answers. In fact, the very talented student who correctly answers all but one of the questions on the verbal part of the SAT loses *fifty* scale points for that one error, earning a score of 750 rather than 800.

What does this mean for the interpretation of aggregate SAT scores? Such aggregates are computed by averaging the individual scale scores of many thousands of students. Since the scale scores summed in those averages come from students representing the entire talent range, it is literally impossible to compute exactly what a difference in aggregate SAT scores means in terms

of average numbers of correct answers. If we assume that most students answer questions correctly, then a decline of 10 points probably means that the students are getting fewer questions correctly. Since the SAT score has declined, it means that the "terrible" decline is perhaps 5 percent in the number of questions answered correctly. In any case, the decline in aggregate SAT scores is not nearly so "massive." Most aggregate SAT scores ceased in the mid-1970s and have not risen since then. Recent annual aggregate SAT scores often amounting to only a few points above the scores of the students of America in 1941. The addition of an additional SAT question would not change the aggregate score.

Third, problems are also created by the use of aggregate SAT scores as averages of individual scores. The value of an individual score when used for this purpose is distorted when aggregated for evaluating districts, or states, and such comparisons are meaningless. The reason is that the aggregate score is only by those high schools in the district. The proportion of students taking the test varies from *less than 10 percent* in Connecticut, Mississippi, North Carolina, and South Carolina to *more than 70 percent* in Connecticut, Hampshire, New Jersey, and New York.

Does this matter? Absolutely. For those students are likely to be those who are trying to get into "the best" colleges. If those of students take the SAT are weak, it is a reflection of their weak high school records. The aggregate score earned by a school, school district, or state is a reflection of the educational quality of that unit. The aggregate score is not a warning that has already been issued by the State Board, which publishes the aggregate score.

Now, let us apply this to the aggregate SAT scores for the nation. The aggregate SAT scores gradually, but since the number of students taking the test has increased, the aggregate scores, therefore, have not

of average numbers of questions answered correctly on the test. But if we assume that most students answered a middling number of questions correctly, then a decline of "from 60 to 90 points" in aggregate SAT scale scores probably means that the average student answered from six to nine fewer questions correctly. Since students had to answer 138 questions in all, this means that the "terrible" decline in SAT scores was in reality a drop of perhaps 5 percent in the number of questions answered correctly. Put this way, the decline in aggregate SAT scores that began about 1963 seems a good deal less "massive." Moreover, the drop that did occur in aggregate SAT scores ceased in the mid-1970s, and there has been no evidence of a decline since then. Recent annual shifts in aggregate national scores have been minute, often amounting to only one or two scale points, which probably means that the students of America have answered, or failed to answer, one or two *tenths* of an additional SAT question.

Third, problems are also generated when one tries to figure out the meaning of *aggregate* SAT scores. The SAT was designed to predict the grade point averages of individual college freshmen and has demonstrated real, if limited, value when used for this purpose. SAT scores were never intended to be aggregated for evaluating the achievements of teachers, schools, school districts, or states, and such scores have *no* validity when used for such evaluations. The reason is that the SAT is a *voluntary* test and is typically taken only by those high school seniors interested in going to college. Moreover, the proportion of students choosing to take the SAT varies sharply across the country. For example, in 1993 the percentage of students taking the SAT varied from *less than 10 percent* in Alabama, Arkansas, Iowa, Kansas, Louisiana, Mississippi, North Dakota, Oklahoma, South Dakota, and Utah to *more than 70 percent* in Connecticut, the District of Columbia, Massachusetts, New Hampshire, New Jersey, New York, and Rhode Island.¹

Does this matter? Absolutely. When only a few students take the SAT, those students are likely to be people with strong high school records, who are trying to get into "the best" colleges. In contrast, when a larger proportion of students take the SAT, that proportion will include more students with weak high school records, who are merely hoping to qualify for some kind of higher education. This means, of course, that the aggregate SAT score earned by a school, school district, or state is not valid for judging the educational quality of that unit unless all, or a *representative* sample, of its students take the test. This may sound like a harsh judgment, but we are only echoing a warning that has already been circulated by the College Entrance Examination Board, which publishes the SAT (see Exhibit 2.2).

Now, let us apply this principle to the problem of interpreting aggregate SAT scores for the nation. Widespread adoption of the SAT came about gradually, but since the mid-1960s, only about half of all eligible high school seniors in the country have taken the test each year.² Aggregate national SAT scores, therefore, have never represented the nation as a whole but have

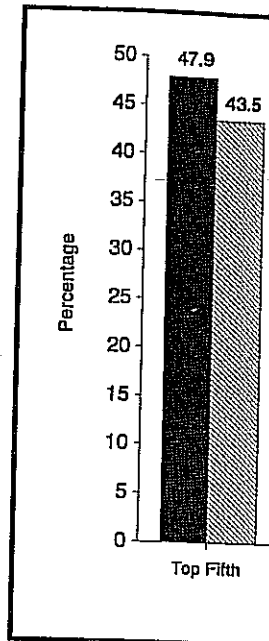
Exhibit 2.3 Percentages

■ EXHIBIT 2.2 Cautions on the Use of Aggregate SAT Scores

As measures of developed verbal and mathematical abilities important for success in college, SAT scores are useful in making decisions about individual students and assessing their academic preparation. Using these scores in aggregate form as a single measure to rank or rate teachers, educational institutions, districts, or states is invalid because it does not include all students. In being incomplete, this use is inherently unfair.

The most significant factor in interpreting SAT scores is the proportion of eligible students taking the exam—the participation rate. Thus, to make useful comparisons of students' performance between states (or among teachers, schools, or districts), a common test given to all students would be required. Because the percentage of SAT takers varies widely, and because the test takers are self-selected, the SAT is inappropriate for this purpose.

—Source: Excerpted from *Guidelines on the Uses of College Board Test Scores and Related Data* (The College Entrance Examination Board, 1988)



—Source: *College Bound Seniors*

always been generated by a self-selected pool of students who wanted to attend college and thought they could improve their chances by taking the test.

Moreover, the composition of students who take the SAT has varied over time. Data on the characteristics of test takers have been distributed by the College Entrance Examination Board since 1976, and these suggest how student composition has changed over the past eighteen years. The accompanying graph (Exhibit 2.3) displays the percentages of students from each of the five high school achievement ranks who chose to take the test in 1976 and 1993. As you can see, more students from the lower achievement ranks have recently opted to take the SAT. This would not matter if students from each achievement rank were equal in their abilities, but this is obviously not the case. Students who earn top high school grades also are much more likely to earn high SAT scores than students who earn mediocre high school grades. This means that even if the ability of schools to educate students remains constant across the nation, aggregate national SAT scores will fall when more students from the lower achievement ranks choose to take the test.

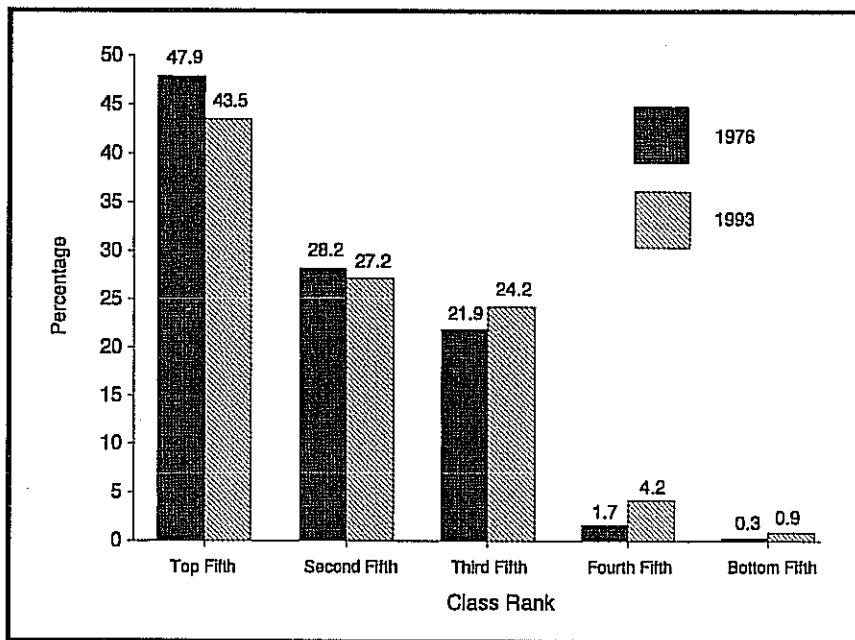
The proportion of students representing minority groups and differing levels of family income has also changed over time, and this has also affected aggregate SAT scores. As we noted above, the SAT was standardized to predict the college grades of students interested in entering college in 1941—at a time when these students were predominantly from white, Anglo-Saxon,

middle- or upper-middle families. We should not be surprised that students tend to have lower SAT scores than fifty years ago. (In fact, Dale Trusheim has recently argued that SAT scores in the 1990s are based on scores that would be expected of impoverished, minority students whose families lack the resources to educate their children. When more students from these families choose to take the test,

Finally, SAT results are also affected by changes in the composition of families. In fact, at present, SAT scores are down by *fifteen points* from what they were in 1976. This means, of course, that when we make decisions, they are also affected by these changes. And it means that aggregate SAT scores are lower from poorer families who choose to take the test.

How does this help us understand the decline in SAT scores during the late 1980s and early 1990s? This was a period of exploding

Exhibit 2.3 Percentage of Students Taking SAT by Class Rank



—Source: *College Bound Seniors* (The College Entrance Examination Board, various dates).

important
ns. about
ng these
s, educa-
clude all
portion of
to make
r among
ts would
and
e for this

Related Data

wanted to
y taking the

varied over
uted by the
st how stu-
accompany-
each of the
n 1976 and
t ranks have
s from each
usly not the
more likely
hool grades.
nts remains
when more
est.

nd differing
also affected
ed to predict
1941—at a
nglo-Saxon,

middle- or upper-middle-class, Protestant homes where English was spoken. We should not be surprised, therefore, to learn that minority and immigrant students tend to have lower scores than the group that first took the SAT fifty years ago. (In fact, so striking is this problem that James Crouse and Dale Trusheim have recently argued that whenever decisions about college admission are based on SAT results, those decisions are *always* biased against impoverished, minority students.³) Even if schools across the nation retain their abilities to educate students, aggregate national SAT scores will also fall when more students from impoverished, minority backgrounds choose to take the test.

Finally, SAT results are also closely tied to the income earned by students' families. In fact, at present the average SAT score earned by students goes down by *fifteen points* for each decrease of \$10,000 in family income.⁴ This means, of course, that whenever colleges use the SAT for making admissions decisions, they are also discriminating against students from poorer homes. And it means that aggregate SAT scores will also fall when more students from poorer families choose to take the test.

How does this help us to understand the small decline in aggregate SAT scores during the late 1960s and early 1970s? If nothing else, this decade was a period of exploding interest in higher education among groups in the

population that had not aspired to college before then. Sharply larger numbers of students from the lower-achievement ranks in high schools, from minority groups, and from poorer families began to take the SAT during those years—and these decisions alone were sufficient to generate the decline in aggregate national SAT scores during these decades. Thus, the brief decline in SAT scores a generation ago provided no information whatever about the performance of American schools but was, instead, a signal that interest in higher education was spreading throughout the nation. Surely this should have been a matter for rejoicing, not alarm.

In larger terms, since the SAT is voluntary and is only taken by roughly one-half of high school seniors across the nation, aggregate national SAT scores will *always* reflect the characteristics of students who choose to take the test. And since those characteristics change over time, aggregate national scores simply *cannot* be used for making valid judgments about the performance of the nation's schools. Despite repeated claims by the critics and alarmist reports in the press, shifts in national *aggregate* SAT scores tell us nothing at all about the performance or problems in American education.

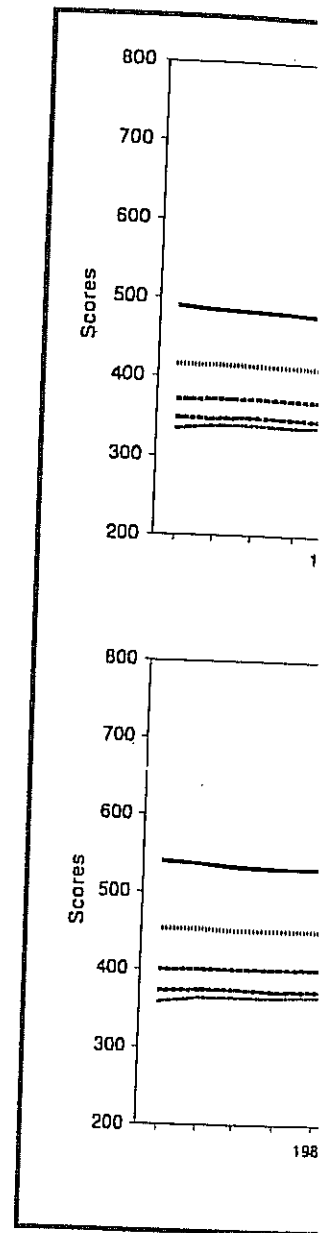
So much for *aggregate* SAT scores. But isn't it possible to discover something about the nation's educational achievements by looking at *disaggregated* SAT data? Yes, it is possible, and it can be done in various ways. For example, we can look at disaggregated SAT scores for students with stronger and weaker high school records, and we can ask, how have these different student groups fared recently in SAT scores?

Exhibit 2.4 displays average scores for the verbal and mathematics subtests of the SAT for students from each high-school rank for the past eighteen years. As this exhibit shows, over this period verbal SAT scores remained quite constant, but scores for the mathematics subtest increased slightly for each of the five achievement ranks. (Granted, these increases were small, but they were consistent. The increase for the top rank was ten points, and for the remaining ranks the increases were thirteen, thirteen, seven, and four points respectively.) Thus, students with both strong and weak achievement records have recently been holding their own on the verbal SAT and doing slightly better in mathematics. Moreover, the top fifth of high school seniors have continued to score at a healthy level on both subtests of the SAT and still earn an aggregate score that exceeds 1000.

In addition, data have been available since 1976 concerning the race and ethnicity of students taking the SAT. As Exhibit 2.5 shows, average SAT scores were nearly constant for white students, but the scores *increased* for every minority group during this period. (The slight decline for white students merely reflects the larger numbers of those students with weaker academic backgrounds who are now taking the test—look again at Exhibit 2.3.)

When one looks at *disaggregated* SAT data, then, one discovers the following: (1) scores for verbal achievement have been holding steady; (2) scores for mathematics achievement have shown modest recent increases; (3) white

Exhibit 2.4 Average Scores



—Source: *College Bound Seniors* (The

ger numbers
om minority
uring those
e decline in
rief decline
er about the
t interest in
this should

by roughly
ational SAT
oose to take
ate national
t the perfor-
e critics and
cores tell us
education.

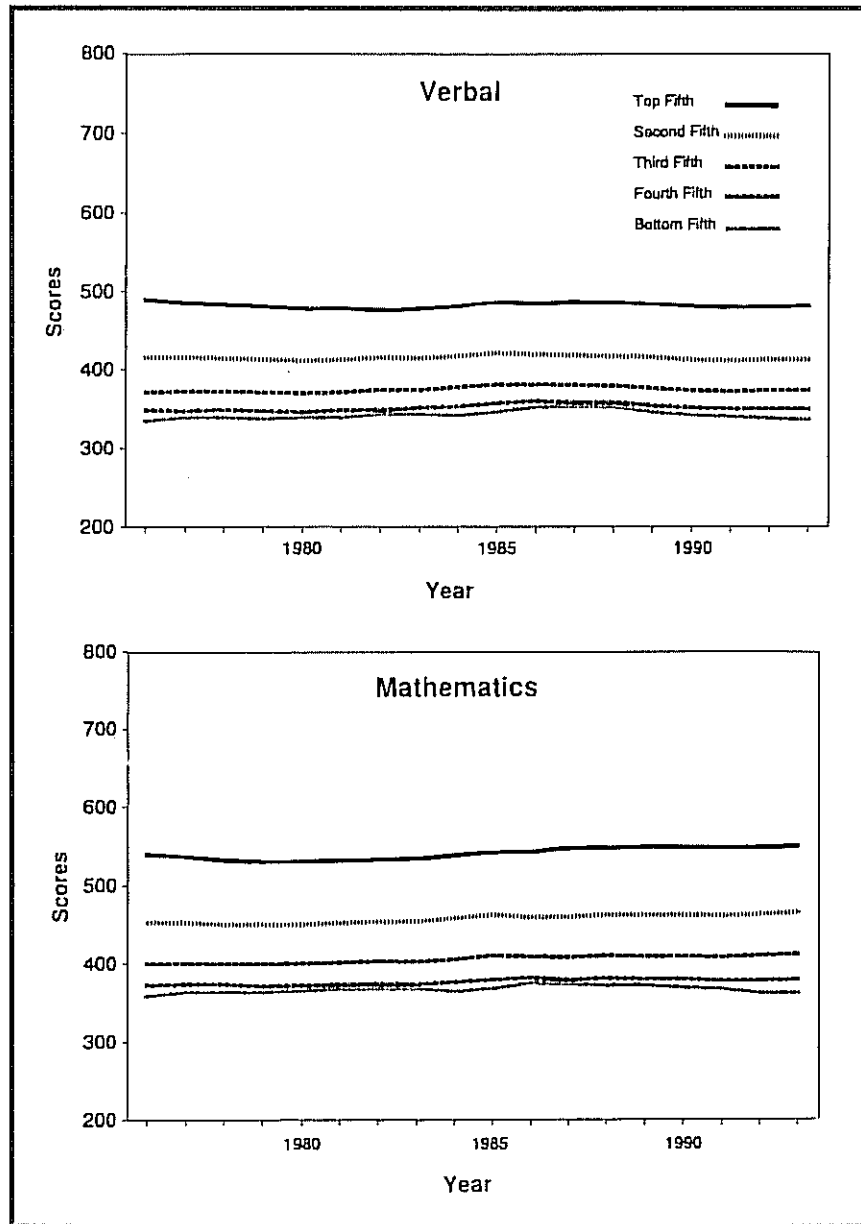
cover some-
disaggregated
For example,
r and weaker
ndent groups

atics subtests
past eighteen
es remained
d slightly for
re small, but
ints, and for
en, and four
achievement
T and doing
chool seniors
the SAT and

the race and
ge SAT scores
used for every
white students
ker academic
bit 2.3.)

ers the follow-
dy; (2) scores
ses; (3) white

Exhibit 2.4 Average SAT Scores by High School Class Rank



—Source: *College Bound Seniors* (The College Entrance Examination Board, various dates).



- Since the SAT is taken by only those students interested in going to college, *aggregate* SAT scores should not be used for judging the performance of schools, school districts, states, or the nation as a whole.
- *Disaggregated* SAT scores suggest that student achievement in the nation has either been steady or has been climbing over the past eighteen years.

So although critics have trumpeted the “alarming” news that aggregate national SAT scores fell during the late 1960s and the early 1970s, this decline indicates nothing about the performance of American schools. Rather, it signals that students from a broader range of backgrounds were then getting interested in college, which should have been cause for celebration, not alarm.

The American College Testing Program. Before turning to other data sources, we should also discuss briefly the other major test that is taken by high school seniors who hope to go to college. The American College Testing Program (ACT) was founded in 1959. Many students now take the twelfth-grade ACT test instead of the SAT, and some critics have claimed that ACT scores also confirm a recent “decline” in the performance of American schools. At first, this claim seems to make sense, since the ACT test is similar to the SAT in several ways. Like the SAT, it covers more than one subject (before 1989, English, Mathematics, Social Studies, and Natural Science). Like the SAT, the ACT test is voluntary. And though it uses a different scaling procedure, like the SAT it uses a fixed process, with obscure rules, for converting number of right answers to scaled scores. Shouldn’t we, then, also look at ACT test scores over the years to check whether they confirm a national achievement decline?

Actually, we shouldn’t do this at all. The reason is that the ACT test does *not* have a fixed set of content topics but, rather, is revised each year by panels of consultants. Each year these experts propose new items to test the *evolving* curricula at American colleges and universities for the subjects tested, and since those curricula change over time, so does the coverage of the ACT test.⁵ This means that average ACT test scores for any given year should *not* be compared with those of other years because the tests they came from were measuring somewhat different things.

Thus, one can safely ignore claims by critics about recent shifts in average ACT test score results. Like the rumors of Mark Twain’s death, the meaningfulness of such shifts (if any) is greatly exaggerated.

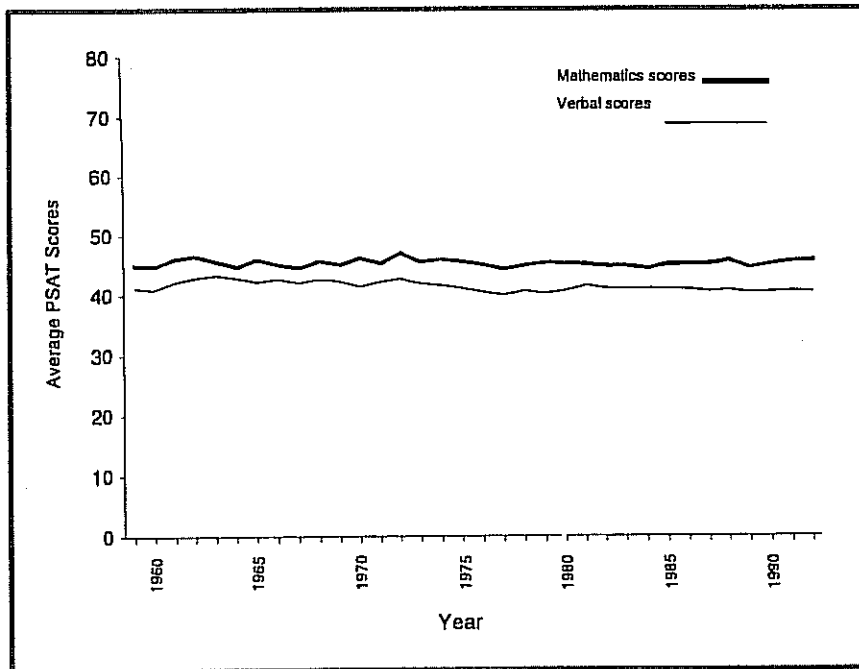
The Preliminary Scholastic Aptitude Test. Given problems with the SAT, it is useful to look also at other test data that may reflect the performance of students in American schools. We turn first to the Preliminary Scholastic

homes are
particular, are
the critics.
ation when
average SAT
nd average
ens should

high
a fixed
mance
northeastern
1941.

large shifts
SAT test

Exhibit 2.6 Average PSAT Mathematics and Verbal Scores



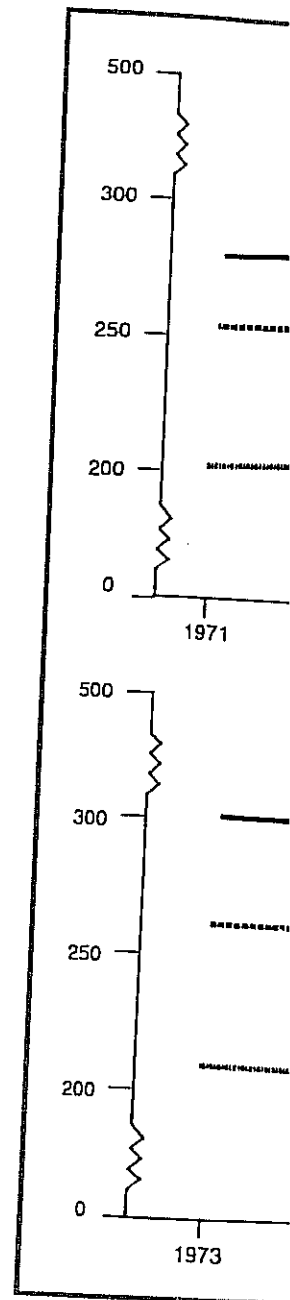
—Sources: Solomon (1983, pp. 24 & 27) and Educational Testing Service (various dates).

Aptitude Test (PSAT). A short version of the SAT, this test is designed to try out new questions and is given annually to national samples of high school juniors. Since the PSAT is given to *representative* samples, rather than to volunteers only, aggregate PSAT scores are more useful than aggregate SAT scores for assessing national achievement.

PSAT data are available for each year since 1959. As Exhibit 2.6 shows, they do *not* suggest any fall in the performance of American schools. In the words of Gerald Bracey, "The lines on a graph of average student scores on the PSAT are as flat as the surface of a frozen lake. Nowhere is there any hint of a decline."⁶ In other words, evidence from the PSAT provides *no* support for the myth of a decline in the academic achievement of American students.

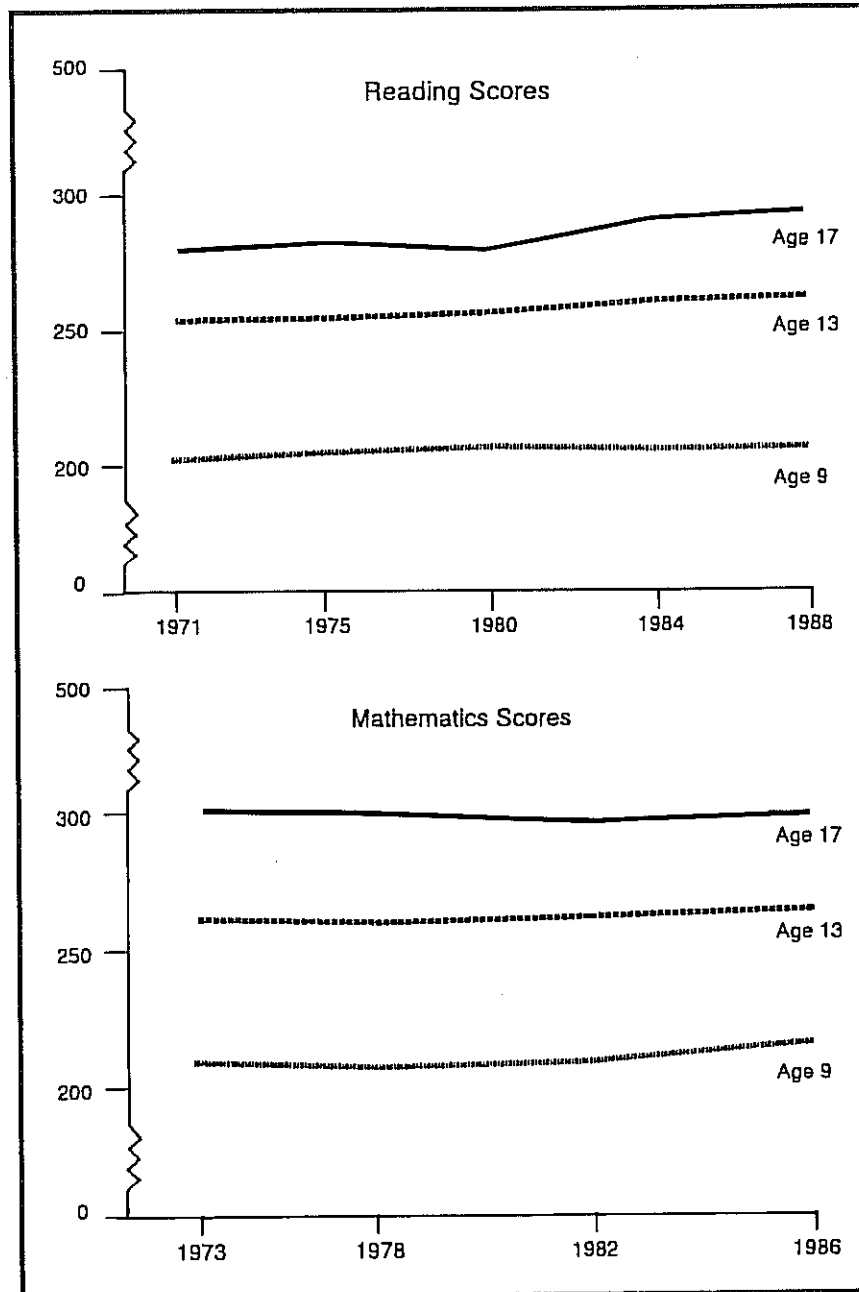
The National Assessment of Educational Progress. An even better source of data about academic achievement comes from the National Assessment of Educational Progress (NAEP). This testing program is conducted by the National Center for Education Statistics of the U.S. Department of Education and is designed to be "The nation's report card" when it comes to assessing educational progress. NAEP tests are given to national samples of

Exhibit 2.7 Nation



—Source: Accelerating Academic A

Exhibit 2.7 National Trends In Proficiency by Year (NAEP Data)



—Source: *Accelerating Academic Achievement* (Educational Testing Service, 1990, p. 32).

designed to
of high school
rather than to
aggregate SAT
it 2.6 shows,
schools. In the
ent scores on
is there any
provides no
of American

An even better
tional Assess-
conducted by
ment of Educa-
n it comes to
nal samples of

students aged nine, thirteen, and seventeen. The students are tested about every two years in mathematics, science, reading, writing, geography, and computer skills. These tests include items that assess reasoning ability and are not intended to be mere tests of memory. What have these tests shown over recent years with regard to student achievement?

In general, the NAEP tests have shown very little change over the past two decades. Exhibit 2.7 shows that average NAEP scores earned by students across the nation in reading and mathematics for various years between the early 1970s and late 1980s have hardly changed during this period.

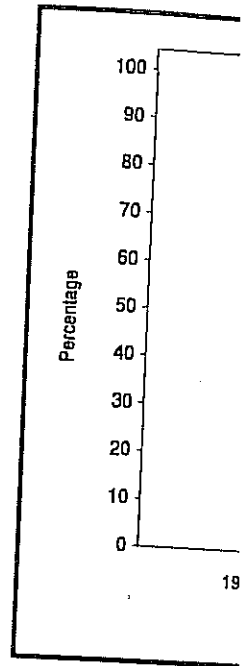
This is not our judgment alone. In 1990 the Educational Testing Service reviewed findings from twenty years of the NAEP and concluded, "there have been various declines and improvements from assessment to assessment, but over the long term, achievement levels are quite stable."⁷ Even such prominent critics of the schools as former Secretary of Education Lamar Alexander and former Assistant Secretary Diane Ravitch were forced to agree with this judgment. Following release of 1991 NAEP scores, Secretary Alexander opined that "today's children seem to know about as much math and about as much science and read about as well as their parents did at that age about 20 years ago"; and Assistant Secretary Ravitch asserted that "the achievement trend lines are essentially flat over the last 20 years."⁸

These general findings hide, however, some interesting NAEP results for subgroups. For example, some NAEP data confirm the recent growth in achievement for minority groups that are reflected in SAT scores. As Exhibit 2.8 shows, the NAEP data indicate that white students have recently held their own in mathematics and that black and Hispanic students have gained significantly.

A second NAEP result suggests a new effect that is difficult to confirm with data from the SAT. Exhibit 2.9 displays average NAEP scores for reading proficiency earned, respectively, by students in Advantaged Metro/Suburban, Rural, and Disadvantaged/Metro schools for various years from the early 1970s through the late 1980s. Over this period, students' scores from advantaged schools were relatively unchanged, while the achievement of rural and disadvantaged/metro schools showed small gains.

(We have reprinted these two graphs from *The Sandia Report*, a major review of evidence concerning the performance of American schools that was originally prepared in 1990 by officials of the Sandia National Laboratories, a branch of the Department of Energy.⁹ Unlike so many recent works that have discussed American education, *The Sandia Report* actually looked at evidence. Not surprisingly, its findings contradicted the erroneous claims that were then being made by education critics in the Bush administration, and as a result the report was suppressed until George Bush was no longer in office. Chapter 4 provides details of this sad story.)

In brief then, evidence from the NAEP also does *not* confirm the myth of a recent decline in American student achievement. Instead, it indicates a

Exhibit 2.8 Perc
17-Ye

—Source: *The Sandia Report*

general pattern of s
achievement among
taged" backgrounds.

Interpretive "Sp
stopped the critics, h
did not confirm the
critics began to shift t
they were certainly no
der, stable NAEP scor
Diane Ravitch declare
enough anymore."¹⁰ A
"students' current ach
below those that migh
English, mathematics,
tional Center for Edu
many students appear
mathematics understar
for college work."¹²

ted about
aphy, and
bility and
sts shown

e past two
y students
etween the
d.

ng Service
there have
ment, but
ch promi-
Alexander
e with this
der opined
ut as much
ut 20 years
ment trend

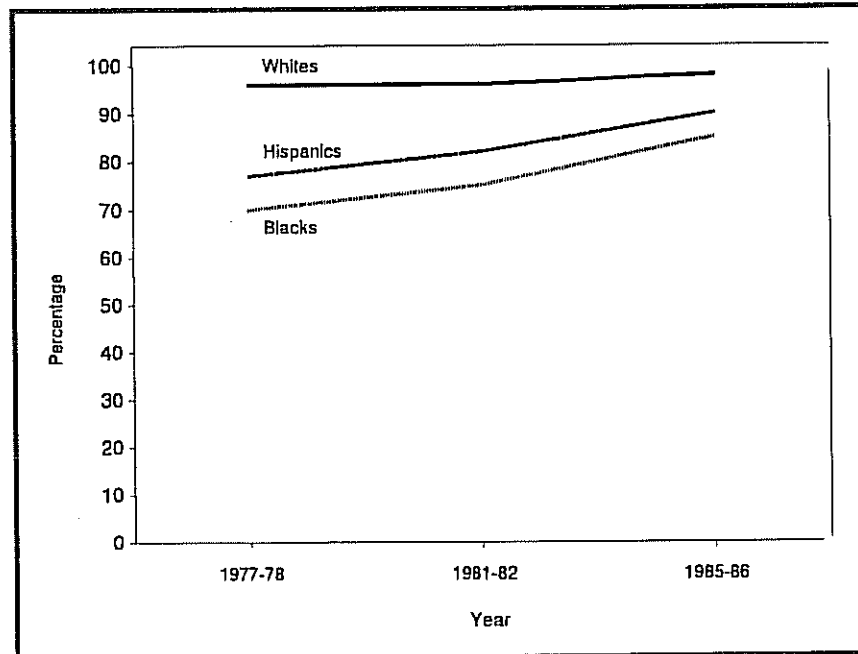
results for
growth in
As Exhibit
cently held
ave gained

to confirm
for reading
/Suburban,
n the early
rom advan-
of rural and

rt, a major
ols that was
laboratories,
works that
y looked at
claims that
ration, and
o longer in

the myth of
indicates a

Exhibit 2.8 Percentage of Students with Basic Math Proficiency
17-Year-Olds (NAEP Data)

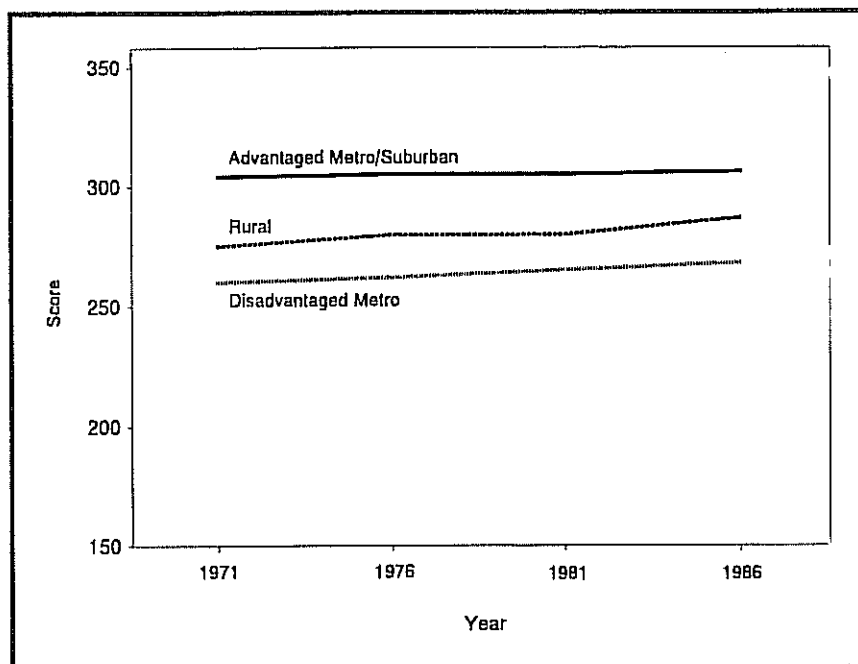


—Source: *The Sandia Report* (Carson et al., 1991, p. 35).

general pattern of stable achievement combined with modest growth in achievement among students from minority groups and from “less advantaged” backgrounds.

Interpretive “Spin.” The “embarrassing” facts just discussed have not stopped the critics, however. Indeed, when it became clear that NAEP data did not confirm the mythic decline of American student achievement, the critics began to shift their ground. If achievement scores were not declining, they were certainly not sufficient for the future. According to Lamar Alexander, stable NAEP scores were “not nearly good enough for the 1990s,” and Diane Ravitch declared that, “what was good enough 20 years ago is not good enough anymore.”¹⁰ As far as the Educational Testing Service was concerned, “students’ current achievement levels [as measured by NAEP scores] are far below those that might indicate competency in challenging subject matter in English, mathematics, science, history and geography.”¹¹ And from the National Center for Education Statistics we learned, “these figures show that many students appear to be graduating from high school with little of the mathematics understanding required by the fastest growing occupations or for college work.”¹²

Exhibit 2.9 Reading Proficiency—Community Types
17-Year-Olds (NAEP Data)



—Source: *The Sandia Report* (Carson et al., 1991, p. 37).

Maybe. It is hard to refute such claims except to note that *they cannot be substantiated* because they are based on unanchored perceptions of national need and on predictions that are not necessarily sound. For example, the jobs of the future may actually require fewer, not more, high-level skills. In an unusual display of agreement, dozens of economists have predicted that growth is likely in the *service* sector of employment—and this means more jobs for janitors, limousine drivers, word processors, sales clerks, and the like. We've also seen estimates that the hospitality industry—e.g., tourism—is now employing more people than any other and that the Wal-Mart chain will soon be the largest single employer in America. But most jobs in hospitality and retail sales do not require high-level mathematical skills. So if schools do not prepare everyone to perform high-level mathematics, perhaps it is because students and their teachers are responding sensibly to the looming job market.

Such musings may also be off the mark, however. What is important to point out here is that some critics confuse what education has accomplished with what one might want it to accomplish. As we suggested in Chapter 1, they muddle reality with desire. Although we would enjoy debating what

MYTHS ABOUT

skills America
than the avera
run down Ame
accomplishmen
deteriorating. W
society that is fa
It is time to cele

Recent interp
used by some crit
some items that
plex reasoning a
NAEP tests delib
tests could measu
and gifted student
only an average sc
on the hardest ite
subjects assessed,
inveigh against the
To illustrate, we

the Educational Te

Performance by hi
graduating from hi
this level of achiev
grounded in comp
year-olds assessed i
mathematical proc
simple equations),
defined—a level ch
multi-step problem

Such judgments a
NAEP were structure
its most difficult que
that the test is doing

We are not the on
Accounting Office (GA
ical and nonpartisan e
examined standards st
out that the Bush adn
deficient in mathemat
level on NAEP exams.
international assessmen
and Eleanor Chelimsky

Fewer than 5 percent of
achievement. . . . For th

skills America might need in the next decade, we claim no more prescience than the average seer. We have little sympathy, however, for critics who run down America's schools for their putative failures when the ongoing accomplishments of those schools are manifest and the society they serve is deteriorating. When school achievements are steady or even improve in a society that is falling apart, we think that educators have pulled off a miracle. It is time to celebrate the public schools of the nation, not to blame them.

Recent interpretations of NAEP data also illustrate another "spin" tactic used by some critics. Like other achievement batteries, the NAEP tests include some items that are easy and others that are very difficult, that require complex reasoning and problem-solving skills. Indeed, those who created the NAEP tests deliberately included both easy and difficult items so that the tests could measure student achievement across the full range of nongifted and gifted students. But this means, of course, that the average student earns only an average score on such tests. *Never* has the average student done well on the hardest items in NAEP tests of mathematics, science, and the other subjects assessed, and this fact offers wonderful opportunities for critics to inveigh against the schools for failing to educate students adequately.

To illustrate, we quote from the recent review of NAEP data prepared by the Educational Testing Service.

Performance by high school students was even more unsettling. Although students graduating from high school seem to be able to add, subtract, multiply, and divide, this level of achievement is hardly in the spirit of our country's goal, which is grounded in competency with challenging subject matter. Only half of the 17-year-olds assessed in 1986 demonstrated a grasp of even moderately challenging mathematical procedures and reasoning (i.e., decimals, fractions, and percents; simple equations), and only 6 percent reached the highest level of proficiency defined—a level characterized by a high rate of success on questions measuring multi-step problem solving and algebra.¹³

Such judgments are nonsense. Like all well-designed tests, those of the NAEP were structured so that only the truly talented student could answer its most difficult questions, and the results cited by the ETS only indicate that the test is doing its job.

We are not the only ones to notice this flaw in reasoning. The General Accounting Office (GAO) is a watchdog agency designed to provide nonpolitical and nonpartisan evaluations of government programs. In 1992 the GAO examined standards set by the NAEP governing board. The GAO pointed out that the Bush administration had claimed that America's students are deficient in mathematics because fewer than 5 percent reach the "advanced" level on NAEP exams. But that same "advanced" level was also used in an international assessment of educational progress in many different countries, and Eleanor Chelimsky, head of the GAO noted,

Fewer than 5 percent of the 9-year-olds in any nation . . . demonstrated advanced achievement. . . . For the 13-year-olds, 10 percent of the students in Taiwan and

at least 5 percent of [a restricted sample from] China . . . met this standard; in no other nation tested did as many as 5 percent meet the advanced threshold. This comparison indicates that the advanced level is extreme even by world class standards.¹⁴

As professors, we have made up hundreds of tests, and we have learned how easy it is to design tests on which nobody can achieve a high score. It is much easier, actually, to assess what people do *not* know than what they do know because the former far exceeds the latter. And being aware of this, we suspect that some of the NAEP effort focuses more on lack of knowledge than on educational progress.

In their zeal to bash American education some critics have also confused test performance with problem solving in the real world. Until recently, major testing organizations in the United States have largely relied on multiple-choice test items to assess what people know and can do—and such items have traditionally appeared in the SAT, NAEP, and related instruments. Testing of this kind simply does not assess the level of real-world thinking possessed by our young people. It confuses in-school problem solving with out-of-school problem solving, two different processes.

We know of one student who is considered quite ignorant in school, particularly in mathematics, but who has recently hired an investment counselor to handle the hundreds of thousands of dollars he made from illegal financial operations when he was a teenager. We do not condone his activities but point out that “school smarts” and “real-world smarts” are not the same thing.

Research supports this claim. Studies have appeared concerning second graders who sell chewing gum on the streets of lesser-developed countries.¹⁵ These children estimate their market, determine markup, factor in inflation rates, determine sales prices, compute discount rates for big purchases, and make change for big bills. But the same children fail in standard tests because they do not know how to solve the in-school problems with approved in-school algorithms. Students with “street smart” mathematical skills often do poorly with some NAEP questions.

Let us look at a sample NAEP item designed to assess competency in mathematics:

Suppose you have 10 coins and have at least one each of a quarter, a dime, a nickel, and a penny. What is the LEAST amount of money you could have?
a) 41 cents b) 47 cents c) 50 cents d) 82 cents.¹⁶

Because most students find it difficult to answer such questions, critics conclude that they will have difficulty with real-world mathematics. But is this reasonable? In the real world, who actually cares about the *least* amount of money one might have from a set of coins? The young people whom we know care more about how coins add up, and those who have “street smarts” have learned to add their coins well.

Why are some adults so adults simply fail to acknowledge to do, and instead focus on arbitrary tasks. To illustrate released, they showed growth of the test. One would Instead, the National Center mathematical skills of our nation either on-the-job demands mathematical literacy.”¹⁷ We would engage in such education measure real-world skills and in the jobs of the future.

Commercial Tests of

ment provide yet another source in America. Each year, millions that are prepared by commercial state of New York, a student taken twenty-nine state-mandated grade.¹⁸ Local districts may irational data are available to examine have it, recently been “failing

These commercial tests are widely known. According to M. Sanders, each year student For example, on average, student ranks for reading and 2.04 percent California Achievement Test (C Test of Basic Skills (ITBS) show oper to write that “composite high in nearly all test areas.”²⁰

country was informed in *A N* declined in virtually every curriculum

The same rising trends were Test (SAT), the Metropolitan Academic Tests of Basic Skills (CTBS). scored higher on both reading and tests. Why then is this effect not be that all these tests are recalibrated is done, the test developers take typical student again scores at t assessed by the tests. In other words achievement gains that students have

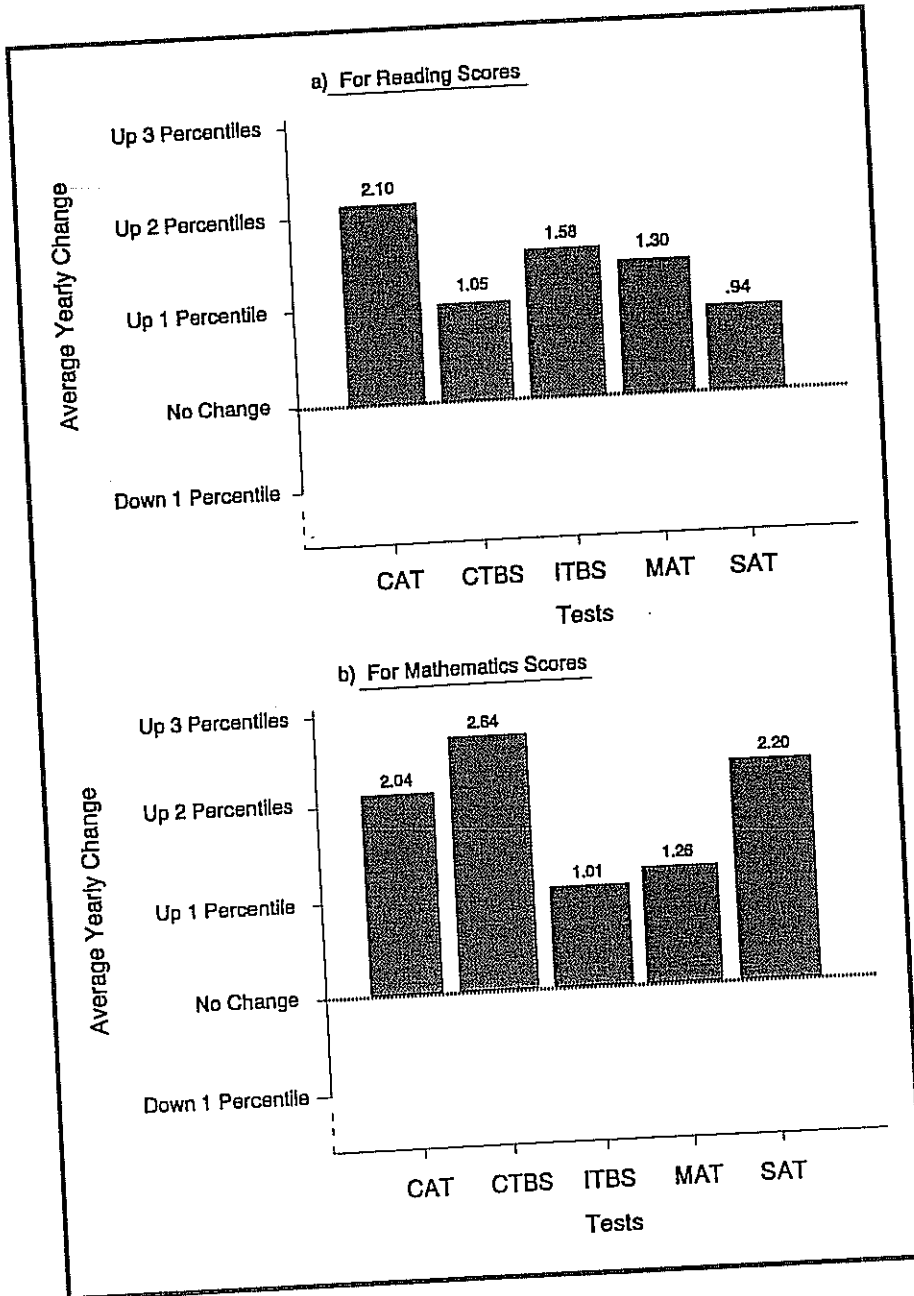
Why are some adults so eager to find young people dysfunctional? Some adults simply fail to acknowledge what our young actually know and are able to do, and instead focus on what youths do *not* know and how they fail at arbitrary tasks. To illustrate, when the 1990 NAEP data on mathematics were released, they showed growth in average scores over *every* previous administration of the test. One would have thought this an occasion for rejoicing. Instead, the National Center for Education Statistics intoned, "The mathematical skills of our nation's children are generally insufficient to cope with either on-the-job demands for problem solving or college expectations for mathematical literacy."¹⁷ We think it's appalling that a government agency would engage in such education bashing on the basis of tests that do not measure real-world skills and have unknown abilities to predict performance in the jobs of the future.

Commercial Tests of Achievement. Commercial tests of achievement provide yet another source of evidence concerning school performance in America. Each year, millions of students all over the country take tests that are prepared by commercial firms. One informant estimates that in the state of New York, a student in a college bound, academic track will have taken twenty-nine state-mandated tests between kindergarten and the twelfth grade.¹⁸ Local districts may impose many more assessments. So lots of additional data are available to examine whether schools have, as the critics would have it, recently been "failing."

These commercial tests are associated with an interesting effect that is not widely known. According to Robert Linn, M. Elizabeth Graue, and Nancy M. Sanders, each year students tend to score higher on these various tests.¹⁹ For example, on average, students in the 1980s gained roughly 2.10 percentile ranks for reading and 2.04 percentile ranks for mathematics, per year, on the California Achievement Test (CAT). These are large gains. Similarly, the Iowa Test of Basic Skills (ITBS) showed large annual gains, leading the test developer to write that "composite achievement in 1984-85 was at an all-time high in nearly all test areas."²⁰ (This statement appeared one year after the country was informed in *A Nation at Risk* that student achievement had declined in virtually every curriculum area!)

The same rising trends were also exhibited by the Stanford Achievement Test (SAT), the Metropolitan Achievement Test (MAT), and the Comprehensive Tests of Basic Skills (CTBS). As shown in Exhibit 2.10, each year students scored higher on both reading and mathematics on all of these commercial tests. Why then is this effect not more widely known? The answer seems to be that all these tests are recalibrated about every seven years, and when this is done, the test developers take pains to make certain that in all cases the typical student again scores at the fiftieth percentile rank for each subject assessed by the tests. In other words, whenever the tests are recalibrated, the achievement gains that students had earned over the past few years are wiped

Exhibit 2.10 Average Yearly Changes in Median Percentile Rank for Well-Known Commercial Achievement Tests



—Source: Linn, Graue, & Sanders (1990, p. 12).

out in the process. So, if cor of them would show that t substantially.

Given these data, we we writing in *The Nation's Busi* average achievement of hig now lower than it was 26 Kilpatrick, like so many of t the facts. In contrast, the ev dent achievement has been

Tests of History and Li
 a specialized study that ha attack public education. D and Diane Ravitch each helk Research and Improvemen 1980s, using NAEP techni school students knew about gloomy book, *What Our I* knowledge was shockingly a barrage of similar argum E. D. Hirsch in *Cultural Lit Mind*, and William Bennet reported these charges as if

Regarding Ravitch and I tory or literature, it suffic designed their test so that : then complained when the : ple, less than 50 percent of

- what made Samuel G famous;
- what the Seneca Falls
- what the issue of “nu
- where Eudora Welty ;
- what John Bunyan’s l
- who wrote, “Things f

(Ask yourself whether *you* evidence certainly reveals l a lack makes American y suggest, or whether it is r

out in the process. So, if commercial tests were not recalibrated, virtually all of them would show that today's students are out-achieving their parents substantially.

Given these data, we wonder where the columnist James J. Kilpatrick, writing in *The Nation's Business*, got his information when he asserted, "The average achievement of high school students on most standardized tests is now lower than it was 26 years ago when Sputnik was launched."²¹ Mr. Kilpatrick, like so many of the critics, seems to have been unencumbered by the facts. In contrast, the evidence from commercial tests indicates that student achievement has been steadily improving over the years.

Tests of History and Literature. Finally, we offer evidence concerning a specialized study that has often been cited by the critics in their zeal to attack public education. During the Reagan and Bush years, Chester Finn and Diane Ravitch each held the post of Director of the Office of Educational Research and Improvement. Together they ran a special study in the mid-1980s, using NAEP techniques and test-item types, to assess how much high school students knew about history and literature. In 1987 they released a gloomy book, *What Our 17-Year-Olds Know*, which claimed that students' knowledge was shockingly thin.²² Their conclusions were, of course, part of a barrage of similar arguments being made at the time by authors such as E. D. Hirsch in *Cultural Literacy*, Allan Bloom in *The Closing of the American Mind*, and William Bennett in *To Reclaim a Legacy*; and the press dutifully reported these charges as if they were gospel.²³

Regarding Ravitch and Finn's claim that today's students know little history or literature, it suffices to point out that these investigators had also designed their test so that it included some truly tough questions, and they then complained when the average student could not answer them. For example, less than 50 percent of 17-year-olds seemed to know

- what made Samuel Gompers, Andrew Carnegie, and Jane Addams famous;
- what the Seneca Falls Declaration concerned;
- what the issue of "nullification" referred to;
- where Eudora Welty and Flannery O'Connor often set their stories;
- what John Bunyan's *Pilgrim's Progress* is about; and
- who wrote, "Things fall apart; the center cannot hold."

(Ask yourself whether *you* know the answers to all of these questions!) Such evidence certainly reveals lack of knowledge, but one wonders whether such a lack makes American youths unfit for citizenship, as Ravitch and Finn suggest, or whether it is relatively minor in a world where knowledge has

exploded and children cannot know everything that historians and classicists would like them to know.

The Ravitch and Finn study was associated with a second charge, namely, that today's students know less history and literature than did students in previous generations. Although Ravitch and Finn did not make this claim, the press interpreted their study to mean that public schools were on the decline. (Perhaps this occurred because Ravitch and Finn suggested that the present generation of seventeen-year-olds are "at risk" because they lack historical and literary knowledge, thus implying that earlier generations might not have been "at risk.") In any case, it is useful to examine evidence concerning this second charge.

In a creative research project, Dale Whittington recently examined students' knowledge of history and social studies across the decades. Whittington carefully sought out tests administered from 1915 until recent years and equated them to the extent possible.²⁴ She then compared content covered, item difficulty, scoring procedures, types of students taking these exams, and so forth. (Among other things, she reports that students seem *never* to have known as much social studies material as the test developers wanted them to know. Like others suffering from Socrates' Syndrome, test developers in each generation have had a tendency to find the next generation wanting. Ravitch and Finn fit well into this sour tradition.)

It is crucial to note, however, that Whittington was able to find forty-three items on the Ravitch and Finn test that corresponded to items given in tests from earlier eras; so claims about the "decline" of student knowledge could be checked. On about a third of these crucial items today's students were less knowledgeable, on about a third they scored equally with students from the past, and on about a third they scored better than students from past generations. So much for the charge that today's students know less history. Whittington concludes,

The perception of decline in the "results" of American education is open to question. Indeed, given the reduced drop-out rate and less elitist composition of the 17-year-old student body today, one could argue that students today know more American history than did their age peers of the past.

Advocates for reform of education and excellence in public schooling should refrain from harkening to a halcyon past (or allowing the perception of a halcyon past) to garner support for their views. Such action . . . is dishonest and unnecessary. Indeed, excellence is a goal that should be advocated on its own merits.²⁵

Amen.

Summary. Standardized tests provide *no evidence whatever* that supports the myth of a recent decline in the school achievement of the average American student. Achievement in mathematics has not declined—nor has that for science, English-language competency, or any other academic subject that

we know of. Moreover, *always* been weak. Indeed, never suggested such a deliberate voluntary test and each means that its aggregate American schools. Instead, ment has either been steady more. And, although test homes have tended to hold have recently shown achievement

What, then, should we claim that student achievement in America is "at risk"? So mind the tale of Chickering when thumped by an actual falling! The sky is falling! A diatribe message to Turkey, sanity was finally restored should we conclude about played key roles in selling and company did not invest early and enthusiastic capital. Chicken Littles but mere

Although evidence does doubt that belief in this repeatedly asserted the myth wish the public schools to it to encourage increased it difficult to abandon the cannot believe that so many United States—could have

But endless repetition the evidence makes it clear been growing in specific, hurried when many measures viewing, overworked and of our children are leading

MYTH . . . Continued Re

Today's select students know are so much slacker intelligence

we know of. Moreover, support for the myth of achievement decline has *always* been weak. Indeed, the two of us know of only *one* test, the SAT, that ever suggested such a decline—and, as readers know by now, the SAT is a voluntary test and each year is taken by differing types of students, which means that its aggregate results are not valid for judging the performance of American schools. Instead, the evidence suggests that average school achievement has either been stable or has increased modestly for a generation or more. And, although top-ranked students and those from “advantaged” homes have tended to hold their ground, those from “less advantaged” homes have recently shown achievement gains.

What, then, should we conclude about the critics’ repeated, hysterical claims that student achievement has declined “massively” and hence that America is “at risk”? Somehow, in this case, the critics’ behaviors bring to mind the tale of Chicken Little. Readers may remember that Chicken Little, when thumped by an acorn, began to rush around declaiming “The sky is falling! The sky is falling!” and managed to sell this alarming and unsubstantiated message to Turkey Lurkey and many other barnyard creatures before sanity was finally restored. But if this was truly a Chicken Little episode, what should we conclude about former Secretary William Bennett and others who played key roles in selling the hysteria? As far as we know, Secretary Bennett and company did not invent the myth of achievement decline but were only early and enthusiastic converts. Bennett and his colleagues were not, then, Chicken Littles but merely Turkey Lurkeys.

Although evidence does *not* support the myth of achievement decline, we doubt that belief in this myth will disappear overnight. Critics who have repeatedly asserted the myth may have difficulty abandoning it. Others, who wish the public schools well, may also endorse the myth in hopes of using it to encourage increased public funding. And citizens of good will may find it difficult to abandon the myth because they have heard it so often and cannot believe that so many important people—including presidents of the United States—could have prattled nonsense for so long.

But endless repetition of a myth does not make it true. On the contrary, the evidence makes it clear that student achievement in America has actually been growing in specific, if modest, ways. Remarkably, this growth has occurred when many measures—particularly those for poverty, violence, TV viewing, overworked and absent parents, and the like—indicate that more of our children are leading difficult lives.

MYTH . . . *College-Student Performance Has Recently Declined in America*

Today’s select students know so much less, are so much more cut off from the tradition, are so much slacker intellectually, that they make their predecessors look like prodigies

of culture. The soil is ever thinner, and I doubt whether it can now sustain the taller growths.

—Allan Bloom (*The Closing of the American Mind*, 1987, p. 51)

Most Americans and their policymakers, concerned about the quality of pre-collegiate education, take heart in the large numbers of Americans who receive associate's and bachelor's degrees every year. The harsh truth is that a significant minority of these graduates enter or reenter the world with little more than the knowledge, competence, and skill we would have expected in a high school graduate scarcely a generation ago.

—Wingspread Group on Higher Education (*An American Imperative*, 1993, p. 2)

Education at the college and university level has also come under attack from the critics, with a lot of the same charges leveled against it that are leveled against schools from kindergarten to high school. Critics of higher education worry about many things: the enrollment of too many "unqualified" students in colleges and universities; reports of decline in academic standards; the appearance of nontraditional students in college courses and their need for remedial work; the rise of courses on black history or feminist concerns, which are reputedly keeping students away from science and the "classics"; the reported promotion of homosexual values or of irresponsible heterosexual relationships on campuses; the number of students who now must work while attending college and therefore may not be able to concentrate on serious academic pursuits; and so forth.

It is not easy to refute these many unanchored worries, and we will not do so directly. We note, however, that if Americans are truly committed to a higher-education system that serves the needs of *many* students, then the fact that those students bring differing strengths and weaknesses to colleges and universities is not very important. What *is* important is that our colleges and universities teach those students well.

Indeed, we shouldn't forget the reaction of some university faculty and graduates to the enormous numbers of veterans who entered colleges after World War II. Some alarmists considered this an onslaught by the unwashed masses, particularly certain ethnic groups who had previously been kept out of higher education—Italians, Eastern Europeans, the Irish, the Jews, and the like. Since the G.I. Bill gave *all* veterans fiscal support to attend college, many children from poor families decided that they too wanted a college education. Allowing many people who did not meet traditional standards an entrance into higher education was declared a disaster by some who already had university degrees and by faculty who had earlier taught only the privileged and "well-mannered."

Not surprisingly, cries about the destruction of higher education were more likely to come from prestigious campuses that served affluent students. For example, the president of the University of Chicago, Robert Maynard Hutchins, suggested that if the G.I. Bill were not amended, it would convert America's colleges and universities into "educational hobo jungles." Instead,

Education money. The C they have eve not want to ke who fail. Even and the vetera will find them.

The remed of national ex benefits has a posed educati and other fee educational in some discrim the administra institutions ha fewer veteran attend who w of Rights sho other ways, m

—Source: Rob

Hutchins propo: pass stiff nation:

In retrospect, were absurd. WI did very well ind were spread in u for widespread required remedi: but in general th

Criticism of c Those who poss others' access to tion system that lege may be disti who are female, universities shou lieve that univers educated and w

■ EXHIBIT 2.11

The THREAT to American Education

Education institutions, as the big-time football racket shows, cannot resist money. The G.I. Bill of Rights gives them a chance to get more money than they have ever dreamed of, and to do it in the name of patriotism. They will not want to keep out unqualified veterans; they will not want to expel those who fail. Even if they should want to, they will not be allowed to, for the public and the veterans' organizations will not stand for it. Colleges and universities will find themselves converted into educational hobo jungles.

The remedy lies in requiring the administrator to ascertain through a series of national examinations whether the veteran applying for the educational benefits has a reasonable chance of succeeding in and profiting by his proposed educational program. The government would pay only half the tuition and other fees of the veteran; the balance would have to be paid by the educational institutions themselves. Thus these institutions would exercise some discrimination in the selection of their students and co-operate with the administrator by admitting only those veterans who could use what the institutions had to offer. These amendments would mean, of course, that fewer veterans would attend colleges and universities; for only those would attend who wanted and could get an education. For these the G.I. Bill of Rights should provide support during their education. Other agencies, in other ways, must tackle and solve the problem of mass unemployment.

—Source: Robert M. Hutchins (1944, p. 21)

Hutchins proposed restricting college entry to those few students who could pass stiff national examinations—see Exhibit 2.11!

In retrospect, we realize that such worries from the privileged campuses were absurd. When post-war veterans flooded into America's colleges, most did very well indeed. Knowledge and values associated with higher education were spread in unprecedented fashion, and Americans' views about the need for widespread higher education changed forever. Some of those veterans required remedial help and special attention to get through college, of course, but in general they were highly successful.

Criticism of colleges and universities seems unlikely to disappear, however. Those who possess an advantage, say a college degree, often like to limit others' access to that advantage, and people who remember a higher education system that once catered largely to the sons of white, Anglo-Saxon privilege may be disturbed when college classrooms include many more students who are female, ethnic, of color, or less affluent. We certainly believe that universities should maintain high standards and teach well. But we also believe that universities should not serve just the sons and daughters of the well educated and well off. And if American universities need to make special

accommodations for those who are motivated to learn but who do not have all the background skills needed, those accommodations should surely be made. That is what education in a democracy is all about.

One charge made against higher education can be assessed by looking at evidence, however. Critics have asserted that those who are now graduating from America's universities know less than earlier graduates, although the data available to assess this charge suggest that it is groundless. Indeed, students now graduating from colleges and universities may know *more* than those who graduated in previous years.

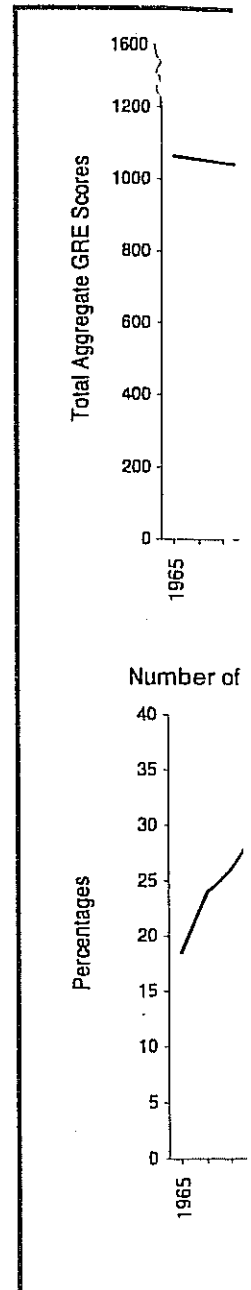
Here we review four sources of evidence that bear on this issue: (1) the Graduate Record Examination, (2) the Graduate Management Admissions Test, (3) the Law School Admission Test, and (4) the Medical College Admissions Test. These four tests are given annually to college seniors who seek entrance to some form of postgraduate or professional education. At least one-third, perhaps one-half, of all college seniors eventually take one or more of these tests, and records from them have been available for many years.

The Graduate Record Examination. The Graduate Record Examination (GRE) is a broad-spectrum test taken by seniors interested in studying for advanced degrees in academic fields. Average GRE scores for all students taking the test are available for each year since 1965—see Exhibit 2.12. Like SAT scores, GRE scores are computed for both verbal and quantitative subtests. As with the SAT, each computation involves obscure rules and yields scores that can range from a low of 200 to a high of 800. Thus, a total GRE score can range from 400 to 1600. And, as with the SAT, percentages of students taking the GRE can differ from year to year; thus, aggregate GRE scores reflect not only student knowledge but also the percentage of students taking the test. Since this is so, we have drawn Exhibit 2.12 so that it provides information about the total aggregate GRE scores and the percentages of students who opted to take the test each year.

As the Exhibit shows, aggregate GRE scores fell slightly from 1965 through the early 1970s, but this was also a period when the percentage of students taking the test *doubled*. Since 1971 the percentage of students taking the GRE has not varied greatly, but *average GRE scores have gradually risen*. What this means is that average total GRE scores are now roughly the same as they were in the 1960s—despite the fact that the percentage of students taking the GRE now is more than twice what it was a generation ago.

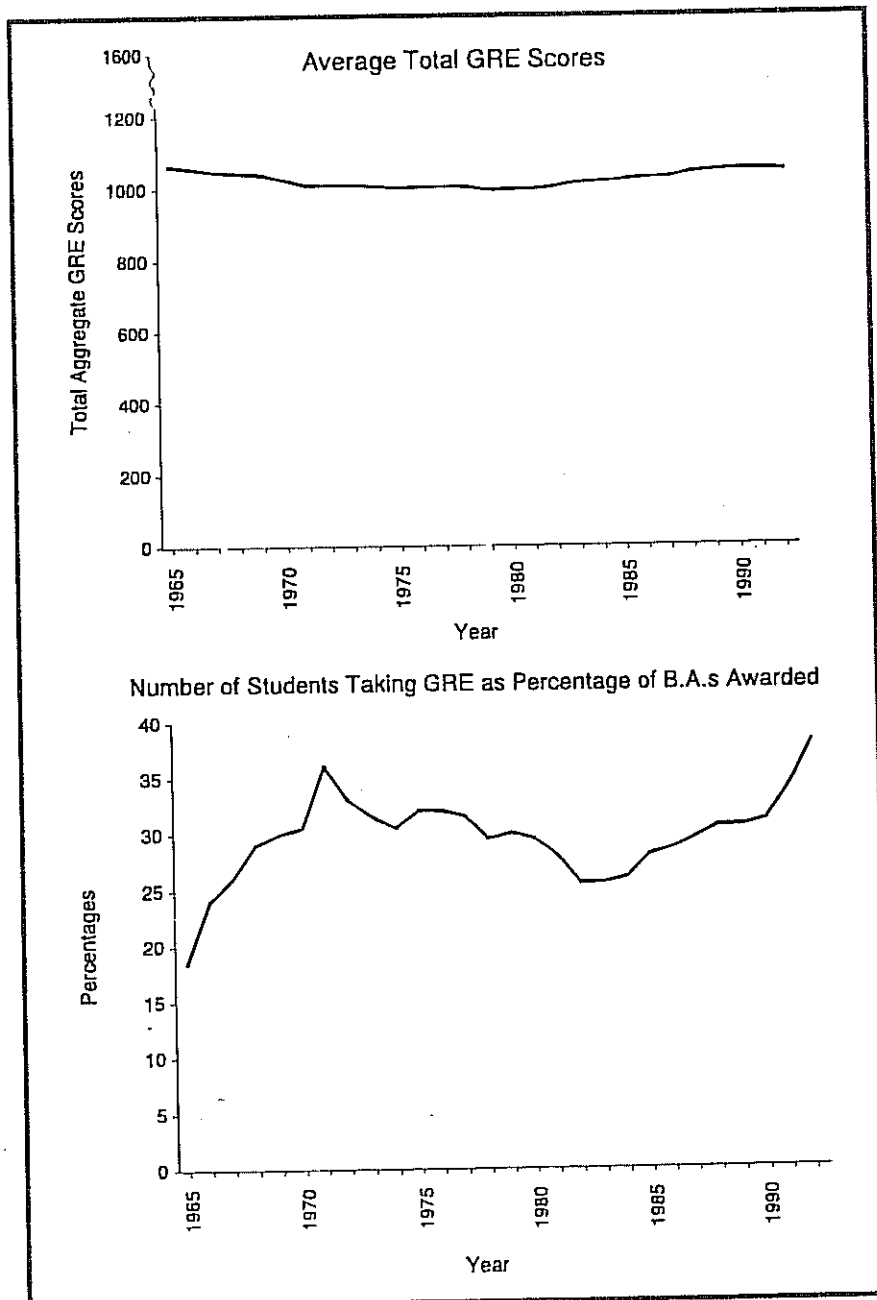
Moreover, the GRE story continues. In 1982 the GRE added a third subtest to its battery concerned with analytic ability. The latter instrument is designed to measure "higher-level" thinking skills; i.e., those associated with reasoning rather than factual knowledge. Aggregate GRE scores for the analytic-ability subscale have also climbed over the past decade and are now about thirty scale points higher than they were in 1982.²⁶

Exhibit 2.12 Average GRE



—Source: *The Condition of E*

Exhibit 2.12 Average GRE Scores and Percentage of Students Taking the GRE



—Source: *The Condition of Education, 1993* (National Center for Education Statistics, 1993b, p. 56).

All of this is indeed good news. Critics have claimed that as more "unqualified" students have been allowed to enroll in America's colleges and universities, the standards of those universities have gone down. GRE evidence suggests that this claim is groundless. True, more students are now enrolled in America's colleges and universities, and a larger proportion of those students are now taking the GRE in the hope of going on to postgraduate education. Despite these greater numbers, college seniors of today are doing just about as well on the GRE as their parents did a generation ago, and evidence from the GRE analytic-ability subscale suggests that their ability to reason is climbing.

Specialized Tests. Results similar to those for the GRE are also available for three specialized tests that are taken by college students who want to enroll in professional schools. The Graduate Management Admissions Test (GMAT) is taken by seniors who want to enter Colleges of Business Administration.²⁷ In 1966 the GMAT was taken by 40,153 students, who earned an average score of 485. Over the next few years, GMAT scores fell slightly, while the numbers of students taking the test soared, but scores have risen appreciably since the early 1970s. In 1992 the GMAT was taken by 231,356 students, who earned an average score of 494. Thus, *five* times as many students are now taking the GMAT, and today's students display just about the same amount of business-relevant knowledge as did those of 25 years ago.

In the case of the Law School Admission Test (LSAT), the average scores earned by college seniors actually *rose* from the early 1960s to the early 1970s—despite the fact that the numbers of students taking the test increased fourfold.²⁸ (For example, average scores for the LSAT were 483 in 1962 and 521 in 1972, while student participation rose from 25,660 to 119,391.) Since then, average LSAT scores have continued to rise, although it is difficult to give comparable figures for more recent test results because different scoring systems were adopted for the test in 1983 and again in 1992.

Over the years, the Medical College Admissions Test (MCAT) has appeared in three different versions, each with altered scoring systems.²⁹ Before 1978 the MCAT consisted of four subtests—measures of Verbal Ability, Quantitative Ability, General Information, and Science Knowledge. Each of these was scored, as are subtests of the SAT, on a scale ranging from 200 to 800. Evidence from this version of the test confirms the same pattern we have seen. Between 1966 and 1975, the number of students taking the MCAT increased from 19,700 to 57,627, but the average scores earned by students increased for three of the four subscales. (Verbal Ability rose from 519 to 541, Quantitative Ability rose from 548 to 583, and Science Knowledge rose from 515 to 567. Only General Information fell—modestly—from 541 to 527.) Since 1978, when the first new scoring system was adopted, MCAT scores have remained relatively constant.

Summary. Gi
GMAT, the LSAT,
continue to charge
has declined. It is
fully tap the types
and we would cert
daunting. But we s
by any great respe
seems likely that i
the critics have be
but are actually fal

But whether thi
dicts a myth about
have just reviewed
have expanded gra
rejoicing. Moreove
know, on average,
years. Americans s
of their colleges an

M

- *Students are*
- *Student inte*
- *Student inte*
school.

I've been watching
as smart as school
you just don't wor

—Emanuel Ber

The common opin
by reason of his su
The children of suc
and ignorant home

—Lewis M. Ter
1981, p. 183)

[The data] make it
of measured intelli
age 17, about 50%
about 30% between

—Benjamin Blc

Summary. Given these four parallel sources of evidence—the GRE, the GMAT, the LSAT, and the MCAT—it is difficult to understand why critics continue to charge that student achievement in American higher education has declined. It is possible, of course, that the tests we have reviewed do not fully tap the types of knowledge that the critics want undergraduates to learn, and we would certainly agree that higher education faces challenges that are daunting. But we suspect that in this case the critics have not been motivated by any great respect for—or even awareness of—the evidence. Instead, it seems likely that in their zeal to express their fears and to motivate reform, the critics have been willing to throw around charges that sound plausible but are actually fabrications.

But whether this explanation is correct or not, the evidence again contradicts a myth about the putative decline of American education. The data we have just reviewed suggest that interest and participation in higher education have expanded greatly in the past generation. This is surely a matter for rejoicing. Moreover, despite this expansion, today's college seniors seem to know, on average, either as much as or a bit more than the seniors of earlier years. Americans should take pride in this evidence of the broad strengths of their colleges and universities.

MYTHS . . . *About Intelligence:*

- *Students are dumber today than they used to be.*
- *Student intelligence is determined only by inheritance.*
- *Student intelligence is largely fixed before students enter school.*

I've been watching and listening to you guys, and I can tell you that none of you are as smart as school kids were when I was in high school. You have it so easy today, you just don't work at learning enough.

—Emanuel Berliner, to his high-school age son, David, and his friends (c. 1953)

The common opinion that the child from a cultured home does better in tests solely by reason of his superior home advantages is an entirely gratuitous assumption. . . . The children of successful and cultured parents test higher than children from wretched and ignorant homes for the simple reason that their heredity is better.

—Lewis M. Terman (*The Measurement of Intelligence*, 1916, p. 115; cited in Gould, 1981, p. 183)

[The data] make it clear that intelligence is a developing function and that the stability of measured intelligence increases with age. . . . In terms of intelligence measured at age 17, about 50% of the development takes place between conception and age 4, about 30% between ages 4 and 8, and about 20% between ages 8 and 17.

—Benjamin Bloom (*Stability and Change in Human Characteristics*, 1964, p. 88)

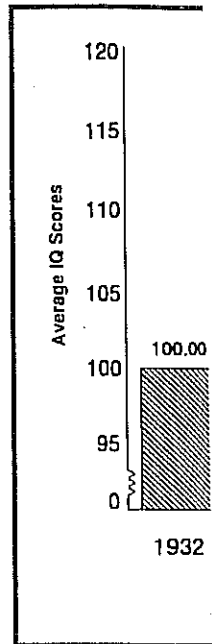
We turn next to three myths about intelligence that are common in America. In fairness, these myths have not often been voiced in recent criticisms of education, but each of these myths has its believers, and that has helped to generate both educational policies and arguments made by the critics of American education. As we shall see, evidence contradicts all three of these myths.

The first myth asserts that young people today are dumber than they used to be. Many untutored Americans, especially those who call in to radio-talk shows, voice the idea that youths of the present generation know less than those of bygone days. Nor are these new charges. In perfect imitation of Socrates, our own parents or their friends sometimes asserted that, when they were young, students were a lot smarter, could think straighter, and had learned more valuable skills than we had. We responded then, of course, by saying that our generation was the brightest that the United States had ever produced. This was merely defensive puffery, but much to our surprise (now that we are the adults), we have learned that our defensive words were actually correct. When we were youths, ours *was* the smartest generation that America had yet spawned. Our delight with this information has been short lived, however, for we have also since learned that the generation of our own children is smarter yet.

The evidence confirming all of this comes from various sources. We begin with a 1987 study by a researcher from New Zealand, James R. Flynn, who analyzed data from fourteen countries.³⁰ When the publishers of the best-known IQ tests—notably the Stanford-Binet and the Wechsler—bring out new editions, which they do every so often, they check to make sure that the new versions generate the same IQ scores as the older tests did. This means that every few years survey data are generated concerning measured IQ scores for the populations of various nations. Exhibit 2.13 presents data from these surveys of IQs among Americans for the years when a new version of either the Stanford-Binet or the Wechsler was developed. As the Exhibit makes clear, whenever a survey was conducted, it found that the average measured intelligence of Americans had increased.

Now this is truly surprising. IQ scores are assumed to be stable throughout a person's adult lifetime, so *growth* in the average measured IQs of Americans presumably means that young people who are coming into the population are brighter, on average, than older people. To be sure, some people believe that these types of IQ tests do a poor job of measuring practical intelligence. It may also be argued that such tests are too narrowly conceived and leave out all sorts of other kinds of intelligence that should be measured—for example, musical ability or visual-spatial capacities. In addition, evidence suggests that standard IQ tests are biased against some groups in the population, particularly minorities and those whose native language is not the language of the test.

Exhibit 2.13 Av
An

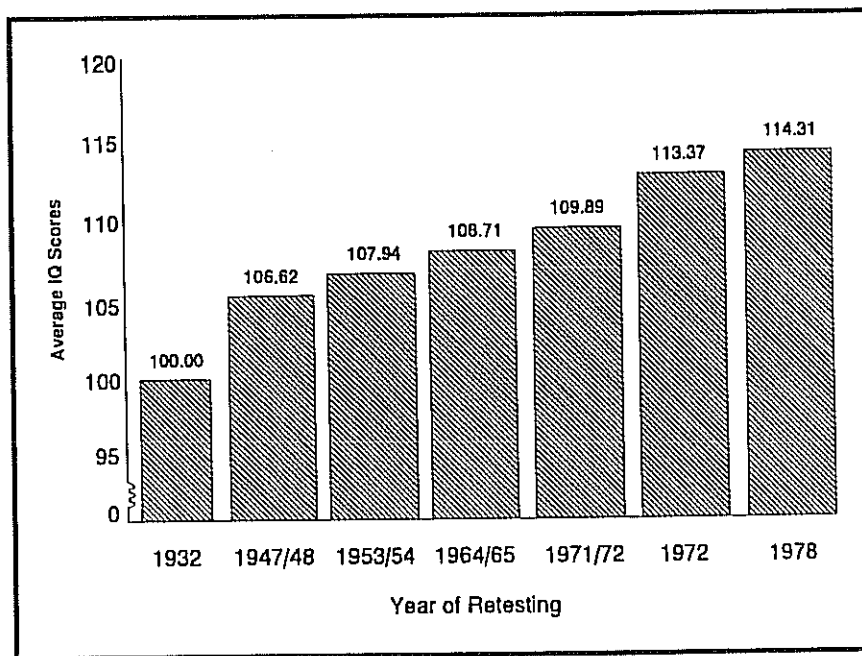


—Source: Flynn (1987, p.

We concur with what is ordinarily shown to relate to kinds of jobs.³¹ A believe in the validity of such tests each year important decisions are generated. So let us

As Exhibit 2.13 shows two to seventy-five Americans for whom to believe that the results. In all of the years, on average, today's youth perform better than their grandparents and the Binet and Wechsler

Exhibit 2.13 Average Wechsler and Stanford-Binet IQ Scores for White Americans Aged 2-75 Years by Year of Retesting



—Source: Flynn (1987, p. 177, Table 7).

We concur with these criticisms, but IQ tests do measure something like what is ordinarily meant by intelligence, and scores on these tests have been shown to relate to success and failure in subsequent education and in various kinds of jobs.³¹ Also, school authorities and school psychologists seem to believe in the validity of these tests since they give hundreds of thousands of such tests each year, spending millions of dollars as they do so, and making important decisions about children on the basis of the scores these tests generate. So let us take Flynn's data seriously.

As Exhibit 2.13 shows, since 1932 the mean IQ for white Americans aged two to seventy-five has risen about .3 points per year. Although it is white Americans for whom the most complete data are available, there is no reason to believe that the pattern is different for other Americans. In fact, Flynn shows that the data from most other industrialized nations show similar results. In all of these nations, people seem to be getting smarter over the years, on average, at least as measured by IQ tests. In the United States, today's youth probably average about 15 IQ points higher than did their grandparents and 7.5 points higher than did their parents on the Stanford-Binet and Wechsler tests. (As with the children in Garrison Keillor's mythical

home town of Lake Wobegone, today's students are, indeed, "all above average.") Or to put this another way, the number of students expected to have IQs of 130 or higher—a typical cut-off point for defining giftedness in many school districts throughout the nation—is now about *seven* times greater than it was for the generation now retiring from leadership positions in the country and often complaining about the poor performance of today's youth. Now *that* is something to contemplate.

Why Are the Young Smarter? With youth throughout the industrialized world showing huge increases in measured IQ, it seems likely that some similar influence is affecting them all. Could it be that young people everywhere are eating better, breathing cleaner air, or drinking cleaner water, and that this has generated the effect? This seems unlikely. The rise in measured intelligence is also correlated with increases in the consumption of fast foods, but we also find it hard to believe that this would cause growth in IQs throughout the world. Instead, we suspect that intelligence, as it is measured by IQ tests, is growing simply because *schools* are doing a better job of promoting the kind of intelligent thinking needed in academic settings. To understand how this has happened, we must briefly discuss IQ tests, heredity, home environment, and the effects of education.

We begin with a simple question: what is meant by *intelligence*? What do people mean when they say that Sally is "smart" and Susan is "dumb"? Such statements usually reflect judgments about a general ability to learn, understand, or solve problems, which is presumed to characterize Sally, Susan, or other people. This definition seems simple, but it also raises two issues. First, unlike height, weight, or eye color, intelligence is a *hypothetical* entity and cannot be observed directly. Thus, intelligence is like the quark or other subatomic particles in physics. It is convenient to think that intelligence is real, but we cannot be certain that it actually exists. Second, intelligence is thought of as a *general* ability, and this means that if Sally is "smart," she can easily learn, understand, or solve problems in various kinds of domains. This last assumption is controversial. Most people are able to accomplish at least some types of tasks—verbal, social, mathematical, mechanical, athletic—more easily than others, and psychologists have debated for years whether to think of intelligence as a single, general ability, or as a set of differing competencies.

In any case, most people today think of intelligence as a single, general ability which, though it cannot be observed, can nevertheless be estimated through appropriate tests. Thus, IQ tests are constructed to assess several types of skills thought to be *generally* useful—memory, reasoning, judgment, and the like. People who take those tests are asked to answer questions or to accomplish tasks measuring these skills, and the tester then computes a single total representing the number of questions answered or tasks the person accomplished correctly.

But how do we level of intelligence compared with the past, and an increase against those other average is given a of 97.5 percent of 130, and so on.) Intelligence generally increases; it is always compared with age. To summarize that presumably rise or to solve problems similar ages from

But what does characteristic, intelligence over time. A French test in 1905.³² Bin which children were he designed a test however, that IQ guaged that intelligence test scores can be compared scores and improve intelligence of an cannot augment. . . mism."³³

Unfortunately, caters for IQ testing as Henry Goddard myth that intelligence largely fixed at conception of Francis Galton, the human genetic the myth of inheritance the forced sterilization "poor genetic stock were an expression

Moreover, if intelligence shows benefit from schooling Burt, also believed

[Intelligence] will both while he is

But how do we know whether such a total indicates a "high" or "low" level of intelligence? To make this last judgment, the total for the person is compared with the average of totals that others have earned on the test in the past, and an IQ score is assigned to indicate how that person stacks up against those others. (By convention, the person whose total matches the past average is given an IQ score of 100. Should that person's total exceed those of 97.5 percent of people from the past, he or she is given an IQ score of 130, and so on.) Since the ability to answer questions and to solve problems generally increases with age, at least among children, the total for each person is always compared with average scores that were earned by people their own age. To summarize, then, IQ tests generate a single score for each person that presumably represents his or her general ability to learn, to understand, or to solve problems when compared with the average abilities of people of similar ages from the past.

But what does the IQ test score imply? Whence cometh this unobservable characteristic, intelligence? Answers to these questions have varied sharply over time. A French scholar, Alfred Binet, created the first standardized IQ test in 1905.³² Binet had been asked to solve a practical problem—to predict which children would have difficulty with standard classroom teaching—and he designed a test to accomplish this task. Binet seems never to have believed, however, that IQ tests measured an inherited, fixed capacity. Rather, he argued that intelligence is malleable and that children who earn low initial IQ-test scores can be given remedial education that will increase their subsequent scores and improve their performance in school. Binet wrote in 1913, "the intelligence of an individual is [not] a fixed quantity, a quantity that one cannot augment. . . . We must protest and react against this brutal pessimism."³³

Unfortunately, such generous arguments were not echoed by early advocates for IQ testing in the United States. Instead, influential Americans such as Henry Goddard, Lewis Terman, and Robert M. Yerkes promulgated the myth that intelligence is strictly a product of inheritance and, as such, is largely fixed at conception.³⁴ In doing so they were responding to the writings of Francis Galton, who had urged that selective breeding be used to improve the human genetic pool; and in the hands of Goddard, Terman, and Yerkes, the myth of inherited intelligence became an excuse for arguments favoring the forced sterilization of "idiots" and immigration laws designed to keep "poor genetic stock" out of the country. Needless to say, these arguments were an expression of xenophobia as well as ethnic and racial bigotry.

Moreover, if intelligence is actually strictly inherited, early measurement of intelligence should enable one to predict completely who will and will not benefit from schooling. To illustrate, the infamous British psychologist, Cyril Burt, also believed in inherited intelligence. In 1947 he wrote,

[Intelligence] will enter into everything the child says, thinks, does or attempts, both while he is at school and later on. . . . If intelligence is innate, the child's

degree of intelligence is permanently limited. No amount of teaching will turn the child who is genuinely defective in general intelligence into a normal pupil.³⁵

Such an overly determined thesis seems absurd today, of course, and the myth that intelligence is determined solely or "largely" by inheritance is espoused today only in works that reveal bigotry and ignorance.

This year's leading candidate for the Bigoted Ignorant Annual Sham (BIAS) award must surely be *The Bell Curve* by Richard Herrnstein and Charles Murray.³⁶ The authors of this widely ballyhooed work argue that intelligence is "largely" inherited, that inheritance alone is sufficient to "explain" average differences in the measured IQs of people who also differ by race, ethnicity, and income, and that federal programs designed to help minorities and impoverished people should be scrapped because they will only encourage such dumb people to breed more rapidly. As we indicated above, these arguments have been around for at least seventy-five years, but in defending them today, Herrnstein and Murray not only had to misinterpret research but were also forced to ignore literally hundreds of studies demonstrating the sizable effects of social environment on measured IQ. These studies are widely known and understood in the large community of social researchers interested in intelligence. Given this widespread knowledge, one wonders why so many media figures have been willing to tout the stale and biased claims of *The Bell Curve*.

Such aberrations aside, by the middle of the twentieth century, most knowledgeable psychologists had come to believe that intelligence reflects both inheritance and the child's home environment. In this they were responding to many studies that had found relations between intelligence and early childhood experiences—particularly parental encouragement, availability of good reading materials, and emphasis on language use in the child's home. For example, in his influential 1964 review, Benjamin Bloom was able to write, "there is little doubt that intelligence development is [also] in part a function of the environment in which the individual lives."³⁷

For decades now, "enlightened" thought has assumed that IQ-test scores will reflect both genetic inheritance and early childhood experience and that all children in the country, therefore, should have an equal right to home and preschool experiences that will help their different intelligences grow. The only trouble with this notion is that it is focused exclusively on *early* childhood experience. In effect, it suggests that whereas intelligence may reflect both heredity and the child's early environment, by the time the child enters school, those influences have played out; thereafter, intelligence is largely fixed. Thus, many people in this country believe that when children enter school they bring with them a relatively fixed capacity for learning, called *intelligence* (which reflects inheritance, home background, or both), and this plays a large part in their subsequent school achievement.

Where did the idea that intelligence was fixed before the early grades come from? Belief in this notion seems to have begun with the early-developmental

theories of psychometrics by Lewis M. Goodenough. This belief was popularized by Benjamin Bloom, who also claims that intelligence appears to be greatest in the early stages of development.³⁸ In retrospect, it is clear that Bloom's theory is faulty reasoning. Bloom's theory is that intelligence is less fixed by age seventeen than it was at age eight but tended to stabilize when it was measured at age eight. This meant that most of the differences in intelligence that have been generated by conditions in the home of children are subjected during the first eight years of life. Head Start program for disadvantaged children is fixed during the first eight years of intervention programs.

Belief in the myth that intelligence is fixed was already well entrenched when the myth was already well entrenched. If it is already well entrenched, it hardly seems surprising that the myth of measured children's IQs in the home is that such tests measure should be based on education, perhaps through traditional or to enrichment programs. So what's wrong with doing such tests with such policies, but we talk about environmental influences are fixed. Instead that the school can also affect intelligence. This implies for educational policy that all children should be thought of as a significant part of the country have an equal right to the same educational intelligences grow. And this means that intelligence influences to some children—because of differences in schools, because they are assigned to different schools they receive substandard curriculum that stunted mental growth.

A Matter of Evidence. Benjamin Bloom indeed modify students' intelligence. Sorel Cahen and Nora Cohen, she asked, as you grow from year to year, how do you mine your achievement at school? How do you mine your measured intelligence?

theories of psychometricians such as E. L. Thorndike and Florence Goodenough. This belief was reinforced by influential reviewers such as Benjamin Bloom, who also concluded that "the effect of the environment . . . appears to be greatest in the early (and more rapid) periods of intelligence development."³⁸ In retrospect, we can see that this conclusion was based on faulty reasoning. Bloom made an assumption that intelligence was more or less fixed by age seventeen. He observed that correlations between early IQ measures and those for seventeen-year-olds varied widely in the early years but tended to stabilize when children entered school, and he concluded that this meant that most of the "development" of intelligence took place before age eight. (It seems not to have occurred to Bloom that the same effect could have been generated by consistencies in the educational experiences to which children are subjected during their school years.) This kind of thinking coincided with the national push for a "war on poverty" and the creation of the Head Start program for disadvantaged preschool children. So Bloom's belief that intelligence is fixed quite early helped provide justification for early-intervention programs.

Belief in the myth that intelligence must be largely fixed at an early age was already well entrenched in America by the early 1970s. And, given this myth, it hardly seems surprising that since then many American schools have measured children's IQs in the early grades assuming that the general ability such tests measure should be assessed in order to plan best for each child's education, perhaps through tracking or by assigning the child either to remedial or to enrichment programs.

So what's wrong with doing this? Actually, there are many things wrong with such policies, but we take on only one issue here. Let's suppose that environmental influences are not confined to the home alone. Suppose instead that the *school* can also affect the intelligence of children. What would this imply for educational policy? It clearly implies that *schools* should also be thought of as a significant source of intelligence and that all children in the country have an equal right to *educational* experiences that will help their intelligences grow. And this means, in turn, that if one denies those experiences to some children—because they are forced to attend underfunded schools, because they are assigned to low-status educational tracks, or because they receive substandard curricula—one is also condemning those children to stunted mental growth.

A Matter of Evidence. But what does the evidence say? Do schools indeed modify students' intelligence? A study by two Israeli psychologists, Sorel Cahen and Nora Cohen, sheds light on this question.³⁹ These researchers asked, as you grow from year to year, does your measured intelligence determine your achievement at school, or does what you achieve in school determine your measured intelligence? In other words, do you have to be intelligent

to profit from schooling (as is widely believed in "enlightened" America), or do you have to have schooling to become intelligent?

Although Cahen and Cohen used complex statistical methods, their findings were straightforward. They found that school achievement was a *major* factor in the prediction of intelligence-test performance. In contrast, measured intelligence was only a *weak* predictor of school achievement. Thus, measured intelligence is strongly influenced by opportunity to learn in school. Over the past fifty years, high quality public education has been offered to larger and larger numbers of students in the industrialized world, and this fact explains why the average person today is measurably smarter than the average person was in the past.

Other evidence supports this idea. Torsten Husén and Albert Tuijnman, distinguished educational researchers from Sweden and Holland, also studied relations between school achievement and measured IQ.⁴⁰ These scholars reexamined data from an older study, originally conducted in Malmö, Sweden, that looked at the IQs of 671 Swedish males over a ten-year period, from childhood to adulthood. Using complex statistical techniques unavailable at the time of the original study, the authors checked whether changes in measured IQ had occurred, and if so what might explain these changes. Their conclusion was unequivocal. Measured IQs *had* changed for many of the persons studied, and those who had experienced more schooling had also grown more in measured intelligence.

Thus, the characteristic that we call *intelligence* is not only dependent on inheritance and home background but is also influenced by schooling. Intelligence during the educative years is not a static and immutable characteristic. It actually appears to be quite dynamic and continues to be affected by environmental factors, particularly by access to high-quality schooling. Husén and Tuijnman concluded,

schools not only confer knowledge and instrumental qualifications but also train and develop students' intellectual capacity. The results [of this study suggest] . . . that IQ as measured by group intelligence tests is not stable but changes significantly between 10 and 20 years of age. . . . [Apparently] schooling co-varies with and produces positive changes in adult IQ.⁴¹

Similar findings were also reported by an American psychologist, Stephen Ceci.⁴² As a result of his research, Ceci concluded that the specific skills measured on intelligence tests and the processes underlying intelligence-test performance are taught and learned *in school*. Ceci also estimated that these influences are substantial. A child could lose as many as six IQ points for each year in which he or she misses high-quality education—from birth onward!

Consider briefly the implication of these studies for typical kindergarten-enrollment policies. If, for example, a school district declares that a child must be five years old by November 30 to enter kindergarten, then a child

born in early December that child may be on November. Suppose, children born in adjacent dates of birth children of roughly the same IQs of those born in October and November of children born in November two groups substantial the children born in November gives children a small over those who start school

In brief, schooling is a major contributor to intelligence. High levels of intelligence are the result of high-quality education for them. Positive lasting consequences

Intelligence, Schooling, and Social Stratification are disturbing. Children born in affluent families that will make the high-quality day care also permits them to pursue and first-rate health care measured intelligence. This is likely to enter school children of poor parents

This sounds bad if we but there is worse. Since it follows that those children most. But since the quality of growth of intelligence is higher for children from rich families families to attend "the best" to private academies or affluent suburbs—and that will the sons and daughters in school funding a private schools are fewer based on student enrollment "buy" intelligence-test performance funding.

born in early December misses out on an entire year of schooling—though that child may be only a day or two younger than children born in late November. Suppose, then, that in this school district we look at the IQs of children born in adjacent months of the same year. With one month difference in dates of birth, we would expect to see little IQ difference among children of roughly the same age—and that is exactly what we find when we examine the IQs of the children born in March and April, for example, or in October and November of the same year. But when we look at the IQs of children born in November and December of the same year, we find these two groups substantially *different* in measured IQ, with the difference favoring the children born in November! Providing schooling at a younger age, then, gives children a small but significant advantage in the form of higher IQs over those who start schooling later.

In brief, schooling matters. Genes and home environment are not the only contributors to intelligence. A society that chooses to nurture and develop high levels of intelligence among its youth must also provide high-quality education for them. Poor schools, like poor home environments, have negative lasting consequences.

Intelligence, Schooling, and Wealth. The implications of this revelation are disturbing. Consider, for example, the effects of wealth on intelligence. Rich parents in America often provide early experiences for their children that will make them smarter. Wealth allows those parents to purchase high-quality day care and to enroll their children in private preschools. It also permits them to purchase instructional toys, encyclopedias, computers, and first-rate health care—all of which are likely to improve a child's measured intelligence. This means that the sons and daughters of wealthy parents are likely to enter school with higher measured intelligence than that of the children of poor parents.

This sounds bad if we believe in the ideal of equal opportunity for children, but there is worse. Since good schooling also leads to gains in intelligence, it follows that those children who attend "the best" schools will also gain the most. But since the quality of schools also varies greatly across the land, the growth of intelligence is not equal in America's schools. And in our country, children from rich families are much more likely than are children from poor families to attend "the best" schools—either because rich children are sent to private academies or because they attend well-funded public schools in affluent suburbs—and this means they will gain more in intelligence than will the sons and daughters of middle- or working-class families. (Such inequities in school funding appear less often in other Western countries, where private schools are fewer and public schools more often receive equal funding based on student enrollment.) Thus, in America we also allow the rich to "buy" intelligence-test points for their children through unequal school funding.

What are the implications for educational policy? In recent years, for example, the Bush and Clinton administrations both urged the states to adopt programs that would provide enriched educational experiences for students who are "truly gifted," and various states have obliged by passing laws setting up such programs. But intelligence tests are widely used to select students who are "truly gifted," and this means that children of the rich are far more likely to participate in such programs than are children of the poor. Many children from poor families could also profit from enrichment programs, but they cannot enter those programs because of the mistake they made at birth—they "chose" to be born to parents who were not wealthy!

This is a problem for our democracy because we continue to preach that all people should have equal opportunities to rise through public schooling. Unfortunately, rising-through-schooling is probably happening *less* frequently in the present decade, in part because so many states have now instituted programs for the "truly gifted." The tests used to select students for these programs measure intelligence or other traits that reflect advantages that only wealthy parents can buy for their children. *Such programs are inherently unfair.*

Other scholars have also made this point, of course, and we will have more to say later about special programs for "gifted" children. Here, we merely repeat that when high scores on intelligence tests are used to select students for enrichment programs, in the United States those programs always confer an unfair advantage on wealthy children.

Our reasoning about wealth and intelligence also suggests that as the number of children living in poverty grows, as it did in the 1980s,⁴³ the continuous rise in intelligence-test scores in this country is likely to stop. The cause for this will not be found in schools but in a society that imposes poverty on growing numbers of its young people. We would be willing to bet, however, that some critics will try to blame public schools for the coming IQ decline.

Summary. What, then, can we conclude from these studies of measured intelligence? First, today's children are smarter, not dumber, than their parents. Furthermore, the parents of today's children were also more intelligent than were their own parents, the grandparents of today's youth.

Second, intelligence is affected not only by inheritance and early childhood experiences; *schooling* also affects IQ test performance. More and better schooling in the U.S. and in other industrialized nations is the most likely reason for those nations' increases in IQ scores. High-quality instructional environments for toddlers, primary-school children, teenagers, and college students all seem to raise scores on IQ tests.

Third, if wealthy people have better access than do poor people to high-quality education, as is the case so often in this country, then some children will do poorly and some will do well on IQ tests because of their parents' wealth rather than because of their genetic makeup or home environment.

And if some students—but if admission programs are systematic—inherently unfair.

Access to schooling, at more intelligent, at quality schooling is others don't; then it will be deprived of innovative businesses to contribute to the

MYTH . . .

International comparison that on 19 academic subjects with other industrialized nations.

—A Nation at Risk

The poor performance of the poor on the front page of the daily newspaper.

—John Murphy

In 1967 a new organization of Educational Achievement compared average mathematics scores of the same age in twelve countries. The results of achievement tests had in so many countries had not finished its comparative studies in science, reading, literature to the surprise of many comparisons. In general, the U.S. lagged behind those of Korea, and Western Europe in last place!

The IEA studies we began to attract wide attention highlighted in *A Nation at Risk*. America faced an educational crisis. The studies indicated serious

And if some educational programs provide advantages for "gifted" students—but if admittance to them requires a high IQ score—then those programs are systematically biased against poor children. Such programs are inherently unfair.

Access to schooling in this country and elsewhere has made people a lot more intelligent, at least as intelligence is normally measured. But if high-quality schooling is rationed in such a way that some have access to it and others don't, then millions of students will be shortchanged, and America will be deprived of the wonderful discoveries, inventions, art and music, innovative businesses, and political leadership those people will never be able to contribute to the country.

MYTH . . . *American Schools Fail in Comparative Studies of Student Achievement*

International comparisons of student achievement, completed a decade ago, reveal that on 19 academic tests American students were never first or second and, in comparison with other industrialized nations, were last seven times.

—*A Nation at Risk* (1983, p. 8)

The poor performance of American schools is now so well known that it makes the front page of the daily newspaper and is a source of public humiliation.

—John Murphy and Jeffrey Schiller (*Transforming America's Schools*, 1992, p. 1)

In 1967 a new organization, the International Association for the Evaluation of Educational Achievement (IEA) brought out a two-volume work that compared average mathematics achievements for secondary students of the same age in twelve countries.⁴⁴ This was the first time that comparable achievement tests had ever been administered to large samples of students in so many countries, and the work attracted a lot of attention. The IEA had not finished its labors, however, for in short order it generated other comparative studies reporting average student achievements for such subjects as science, reading, literature, native-language instruction, and the like. Much to the surprise of many people, the United States generally did *poorly* in these comparisons. In general, American students' average achievement scores lagged behind those of students in comparable countries—particularly Japan, Korea, and Western Europe—and sometimes American students' scores were in last place!

The IEA studies were mainly written for professionals, but eventually they began to attract wider attention. Then in the early 1980s, the studies were highlighted in *A Nation At Risk* and other critical documents claiming that America faced an educational CRISIS. These sources argued that the IEA studies indicated serious deficiencies in American public education and that

those deficiencies were responsible for America's inability to keep up with other nations in industrial competition.

This was heady stuff, of course. As at the Olympics, national pride was involved, and the media rose to the occasion. They began to report that American schools and teachers were failing badly by international standards. Those who wanted to discredit public schools were given new ammunition, columnists provided endless commentary on the "sad" state of American education, and even educators and supporters of public schools began to cringe and to seek ways to explain away the apparent "weaknesses" of American education revealed by comparative research.

Unfortunately, most people who have claimed that American education lags behind education elsewhere have never bothered to *look* at the evidence, and few authors have taken the time to *think* about the assumptions involved when one compares aggregate achievement scores among countries. If one actually looks at and thinks about the comparative evidence, however, one discovers that it does *not* confirm the myth of American educational failure. Indeed, it suggests that in many ways American education stands head and shoulders above education in other countries. Let us, then, look at and think about the evidence of comparative research.

Different Visions. We begin by noting that countries vary greatly in their notions about childhood and how to conduct education. This sounds like a truism, but it is also a fact that generates serious problems for comparative studies of achievement. For example, countries in continental Europe have traditionally used stiff national exams to sort out students at the end of primary or junior-high education, and only those who pass such exams are allowed to enter specialized high schools, *gymnasias*, that prepare them for university entrance. Many Asian countries also use such examinations today, but the United States does not. Does it make any sense, then, to compare the average, national achievements of high school students in mathematics, science, or literature from countries with such disparate systems of education?

This issue becomes crucial when one ponders the unique values reflected in American education. To begin with, Americans think that children should have a wide variety of experiences. Our middle-class neighbors seem to agree that their children should participate in organized sports such as Little League, basketball, and soccer; engage in after-school activities such as piano lessons and dance; watch a good deal of TV; spend weekends in leisure pursuits; have their own cars and begin to date while in high school; and so forth. This means, of course, that many American parents do *not* favor an educational system that assigns vast amounts of homework or that encourages students to become high-achieving drudges. By comparison then, American teenagers probably have more nonacademic interests and a wider knowledge base than do students from countries that stress narrow academic concerns.

As a na
people to g
Thus, we
people are
tions—Jap
growth in t
potential st
—Again, A
be socially
Visitors to
America's y
to promote
on the sub:
achievemen

America
mary studen
play musica
schools. An
are encoura
host of extr
in the conce
unique to th
ing as docto
sity at age 1
than they de
can student
where thoug
academic su
they should,
as well as de

What thi
their school
comparative
and on the t
edge, compa
specific acad
of equivalen
seems yet to
none has yet
independen
graduates or
comparative
at the streng

As a nation, we apparently also believe that it's worthwhile for young people to gain work experience and to learn how to handle their own money. Thus, we promote employment for our teenagers; as a result, our young people are more economically active than is the youth of comparable nations—Japan, for example.⁴⁵ (Given some of the critics' desires to promote growth in the nation's economy, why have they so consistently ignored this potential strength of America's young people?)

Again, Americans like their children to be creative, to be spontaneous, to be socially responsive and friendly, and to challenge unreasonable authority. Visitors to our country often comment with pleasure on these qualities in America's young people. But if school experiences in this country are designed to promote these qualities, it may also be that the schools downplay stress on the subservient conformity that generates high levels of subject-matter achievement in some other countries.

Americans are also profoundly committed to breadth of education. Primary students in our country not only study the three Rs, but they also paint, play musical instruments, debate, and compete in chess tournaments in their schools. American high schools offer a huge range of courses, and students are encouraged to sample these courses as electives and to participate in a host of extracurricular activities. This commitment to breadth shows up also in the concept of a four-year, liberal-arts undergraduate education, a concept unique to the United States; students elsewhere begin their professional training as doctors, lawyers, or licensed psychologists when they *enter* the university at age 18. Our system works well because students stay in school longer than they do in other countries. By comparison then, at any given age American students are likely to be more broadly educated than are students elsewhere though they may not yet have as much detailed knowledge of specific academic subjects. They will acquire this knowledge over time, of course; and they should, on average, end up with a knowledge base that is uniquely broad as well as deep.

What this means is that if Americans are truly interested in learning how their schools stack up comparatively, they should insist that at least *some* comparative studies focus on the values that *Americans* hold for their children and on the unique strengths of *American* schools. To the best of our knowledge, comparative studies to date have all examined student knowledge of specific academic subjects, and nearly all have compared the achievements of equivalent-age junior high or high school students. None of the studies seems yet to have investigated breadth of student interests or knowledge; none has yet examined student creativity, initiative, social responsibility, or independence of thought; and few have studied knowledge among undergraduates or young people who have completed their educations. In fact, comparative studies to date seem almost to have deliberately avoided looking at the strengths of American schools!

Problems with the Studies. The critics and the press seem also to have ignored crucial reasons why comparative evidence can be shaky. Many countries have large ethnic minorities that do not speak the dominant language of the country. Schools serving those minorities often do not conduct education in the dominant language and are excluded from comparative research. In contrast, comparative-study samples from America usually include students from families that *do* and *do not* speak English. Many other countries also operate educational systems in which only selected students are allowed to enter high-status high schools, and comparative data from those countries may be collected only from those high-status schools. In contrast, the typical high school in America serves the full range of students, and comparative evidence from this country usually represents all students attending those schools. Such differences mean that American student-achievement scores may look bad simply because they are gathered from the full range of students in the country, whereas scores from other countries are gathered from biased samples.⁴⁶

Further, countries may differ in the conditions that exist when comparative data are gathered. For example, a Quality Control Observer reported the following conditions for testing in South Korea (a nation famous for its high overall achievement scores):

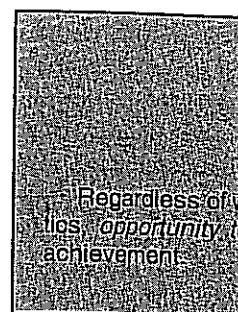
The math teacher . . . calls the names of the 13-year-olds in the room who have been selected as part of the IAEP sample. As each name is called, the student stands at attention at his or her desk until the list is complete. Then, to the supportive and encouraging applause of their colleagues, the chosen ones leave to [take the assessment test].⁴⁷

As Gerald Bracey has noted, these students were apparently taking the test for the honor of their country!⁴⁸ This is surely not likely to occur in America, where our overtested students are accustomed to competing regularly for individual rewards and are likely to view the tests used in comparative studies as an inconvenience offering no obvious personal advantage. Thus, American students may have fewer reasons to perform well in comparative studies.

The point is that in many comparative studies the aggregate scores reported for different countries are not strictly comparable. Such studies are not useless, of course. Indeed, one can learn from them by looking at disaggregated data for subsamples in various countries. But the differences in aggregate national-achievement scores reported by those studies—differences so often relished by hostile critics and the compliant press—may reflect little more than biases in samples and differences in testing procedures.

Opportunity to Learn. A third issue to think about when examining comparative evidence is whether opportunity to learn is the same for different nations. Such opportunities may vary for several reasons. Some countries require that all of their students study key academic subjects—such as ad-

vanced algebra or those subjects are (in America) may lack key subjects are taught until the high school in the aggregate level to identify a single subjects, that factor is so important, and criticize American s Law (see Exhibit 2.



A good example studies of gender and have wondered why roughly equal to that years but suddenly a mystery here. Higher can high schools, and have more opportun

The same process The Second Internati from 1980 to 1982. I (i.e., eighth-graders) found that the aggre lagged behind that of fact was immediately enthusiastically vilifie

Nobody at the tin requiring eighth-grade algebra, whereas such year or two later. Ho at the University of Illi

vanced algebra or specific foreign languages—whereas in other countries those subjects are optional, and some countries (and poorer school districts in America) may lack resources to offer those subjects at all. In some nations key subjects are taught in the early grades, but elsewhere they are delayed until the high school years. Factors such as these make an enormous difference in the aggregate level of student knowledge at a given age. In fact, if we had to identify a single factor that best predicts student knowledge of specific subjects, that factor would certainly be opportunity to learn. This principle is so important, and is so often ignored by those who misuse evidence to criticize American schools, that we restate it here as our *Student Achievement Law* (see Exhibit 2.14).

■ Exhibit 2.14
Berliner and Biddle's
Student Achievement Law

Regardless of what anyone claims about student and school characteristics, *opportunity to learn* is the single most powerful predictor of student achievement.

A good example of the Student Achievement Law at work appears in studies of gender and mathematics achievement. Many good-hearted people have wondered why the mathematics achievement of American girls is roughly equal to that of boys throughout the primary and junior high school years but suddenly drops off in senior high school. Actually, there is no mystery here. Higher-level mathematics courses are optional in most American high schools, and since more boys choose to take those courses, the boys have more opportunity to learn advanced mathematics, and they learn it!

The same process affects comparative studies of mathematics achievement. The Second International Mathematics Study from the IEA was conducted from 1980 to 1982. It looked at the achievement of both thirteen-year-olds (i.e., eighth-graders) and high school seniors. Among other things, the study found that the aggregate achievement of eighth-grade American students lagged behind that of students in many other countries, notably Japan.⁴⁹ This fact was immediately pounced on by critics and by a dutiful press, which enthusiastically vilified American schools for fecklessness.

Nobody at the time seemed to notice that Japanese schools were then requiring eighth-grade students to take mathematics courses that stressed algebra, whereas such courses were typically offered to American students a year or two later. However, this fact was noted by Ian Westbury, a scholar at the University of Illinois who knew the IEA study well. And when Westbury

reexamined the IEA data—looking at the achievement of American students who had and had not taken algebra—he found striking differences in American student achievement. Westbury identified four contrasting math curricula to which eighth-grade American students had been exposed: (1) remedial classes (30 cases); (2) typical nonalgebra classes (174 cases); (3) enriched prealgebra classes (31 cases); and (4) algebra classes (38 cases).⁵⁰ Only the latter two types of classes were comparable to those offered to Japanese students, and when Westbury examined student achievement for those classes, he found records that matched or exceeded those of students in the Japanese schools (see Exhibit 2.15).

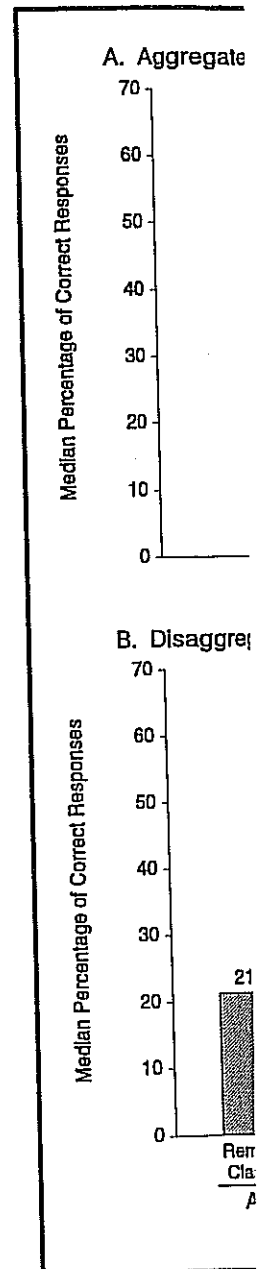
Thus, what the critics had interpreted as a failure of American schools turned out to be merely a reflection of the age at which algebra instruction is typically begun in Japan and America. One may choose to debate whether algebra should or should not be taught at an earlier age, and Americans might want ultimately to change the ways in which instruction is scheduled in our schools. We might wonder, too, what prealgebra is. No other nation seems to have it, and it sounds odd to us. If American schools have students who can do prealgebra, perhaps they should give them algebra instead.

One can debate such curricular issues forever, but what is clear is that the IEA data displayed by Westbury do not suggest that American teachers or schools are deficient compared with those in Japan. Instead, the data reveal merely that the two nations differ in the typical ways in which they structure opportunities to learn. And, if Americans should truly decide that they want their high school students to be “first” in high-level mathematics achievement, all they need do is to fund, offer, and encourage students to take high-level mathematics courses at an early age, in all schools. Is this wise? Ah, that is another question, to which we return later in this book.

Another example of the impact of opportunity to learn also concerns Japanese and American student achievement. From upper elementary through high school, most Japanese students go to private academies called *Juku* schools. Such schools offer additional instruction in academic subjects after regular school hours and on weekends, helping students prepare for entrance exams. Given the time spent in instruction in regular and *Juku* schools combined, the typical Japanese sixteen-year-old will have accrued something like two more years of formal schooling than his or her American counterpart. What is newsworthy if occasionally a comparative study finds that one group of high school students with two more years of instruction happens to outperform another group, which has not had that opportunity to learn?

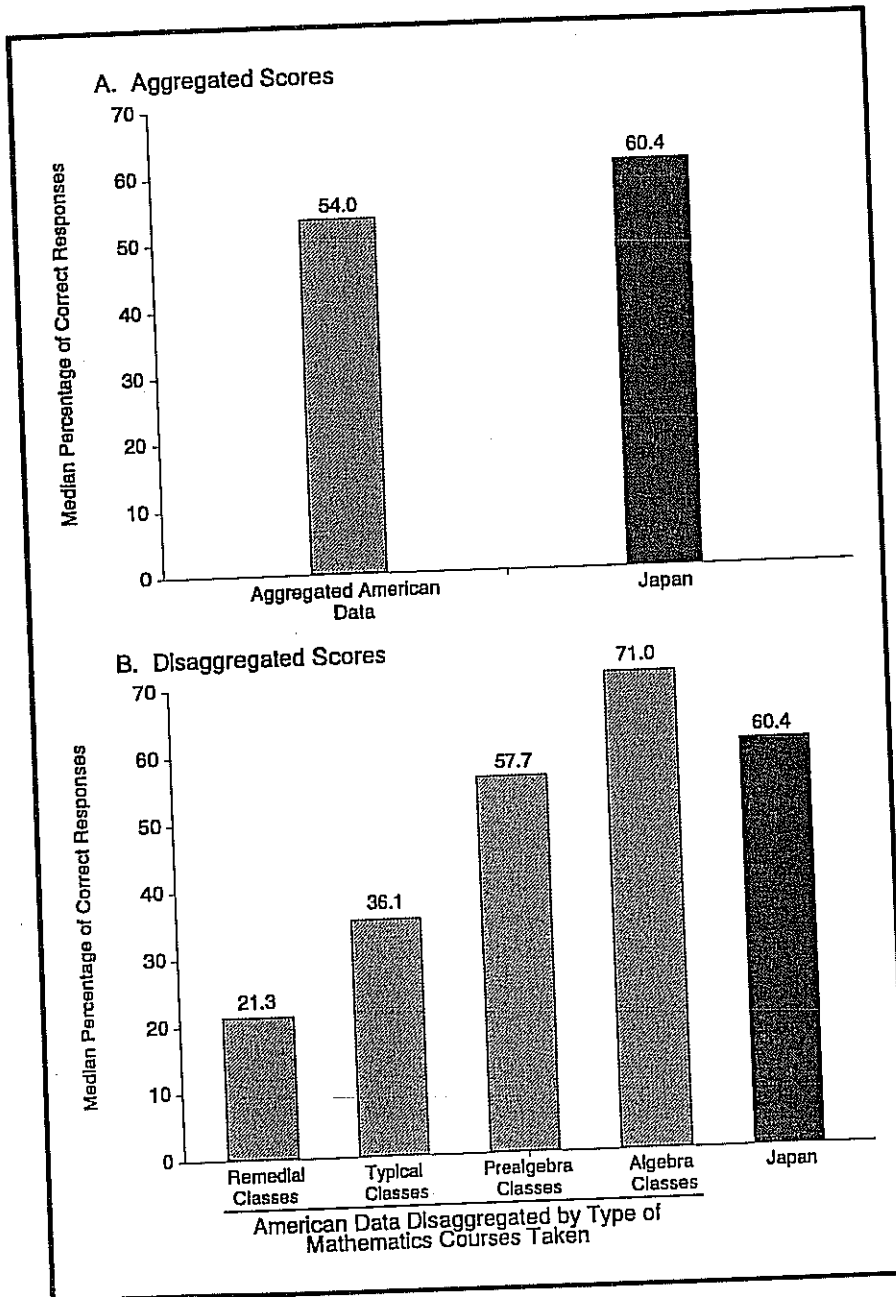
In general then, *aggregate* comparative evidence provides no convincing support for the myth that American students fail in mathematics achievement—or achievement in any other subject—compared with students in other countries. Instead, much of that evidence merely reflects differences in opportunities to learn, and when controls for those opportunities are used

Exhibit 2.15 Japanese
Age 13
of the



—Source: Ian Westbury (1995)

Exhibit 2.15 Japanese and American Achievement Scores for Students Age 13—from The Second International Mathematics Study of the IEA



—Source: Ian Westbury (1992, 1994).

with the data, American students' school achievement looks quite similar to that of students from other countries.

Variability and Test Performance. Yet another issue plagues attempts to compare aggregate scores for student achievement in the U.S. with that of other countries. To state this issue succinctly, the achievement of students from American schools is a *lot* more variable than is student achievement from elsewhere. Thus, any simple attempt to aggregate American students' school achievements may be thought of as akin to reporting the average weight for a group of mice and elephants.

Two serious problems help to generate huge differences in achievement among American students. First, the American population exhibits more extremes of wealth and poverty than does that of comparable nations; the problem of income disparity falls largely on young people; and children from rich and poor families are unlikely to attend the same schools. By comparison, then, the United States has both more extremely rich children and a *much* larger population of poor children than do other Western countries. (By 1980 about one-fifth of all children were living in impoverished families in the United States, and the numbers have gotten worse in subsequent years.⁵¹) This means that vast numbers of poor, ill-clothed, undernourished children from distressed homes with few educational resources now flood our inner-city and poorer rural schools. This extreme problem is simply *not* faced by schools in other Western countries, where rates of childhood poverty are a fraction of those in America⁵² and residential ghettoization is less prevalent. (Readers may think we are exaggerating this problem. The frightening data that confirm it are summarized in Chapter 6.)

Second, Americans tolerate enormous inequities in funding for schools that serve the rich and the poor. Rich people in the United States are able to buy some of the world's finest education for their children, either in private academies or in well-financed, suburban, public-school districts. In contrast, children of the poor are often crowded into miserable rural or inner-city schools whose annual per-student support may be one-fifth or less of that in nearby, suburban public schools. This means, of course, that students in the suburbs have smaller classes, teachers with higher salaries and morale, and more computers and other facilities—whereas schools in slum neighborhoods must contend with dangerous and decaying buildings, gross overcrowding, violence, and inadequate funding for even basic instruction. This problem is also not faced in other Western countries, where equal basic support is normally provided to each student in all public schools.

Together these two problems mean that scholastic achievements will vary far more in the United States than in other countries. To put it baldly, America now has some of the finest, highest-achieving schools in the world—and some of the most miserable, threatened, underfunded educational travesties, which would fail by any achievement standard. This means that any attempt

to aggregat
bound to r

How, th
ing of how
best way to
poverty, pe
ing school
complex sta
that differ i
of *no* comp

We do k
comparing
public scho
other coun
student sup
provided by
recent com
eighth-grad
Dakota, an
the highest-
Korea). In c
issippi) stand

Thus, the
remarkable
it seems fru
of Americar
about stude
to find way
most neglec

The Unri

parative rese
the press see
to the studie
summarized
tional comp
Purta, and I
less than Jap
but that by
to grow in l
senior years
ignored thes

In 1992 t
of reading a

to aggregate the achievements of American schools into a single score is bound to misrepresent the true picture.

How, then, might one use comparative evidence to gain a valid understanding of how American schools stack up against schools elsewhere? Perhaps the best way to do this would be to isolate the independent effects of student poverty, per-student support of the school, and national context when analyzing school achievement data. This analysis could be done either through complex statistical techniques or by disaggregating the achievement of schools that differ in student poverty or per-student support. Unfortunately, we know of *no* comparative study that has provided either of these types of analysis.

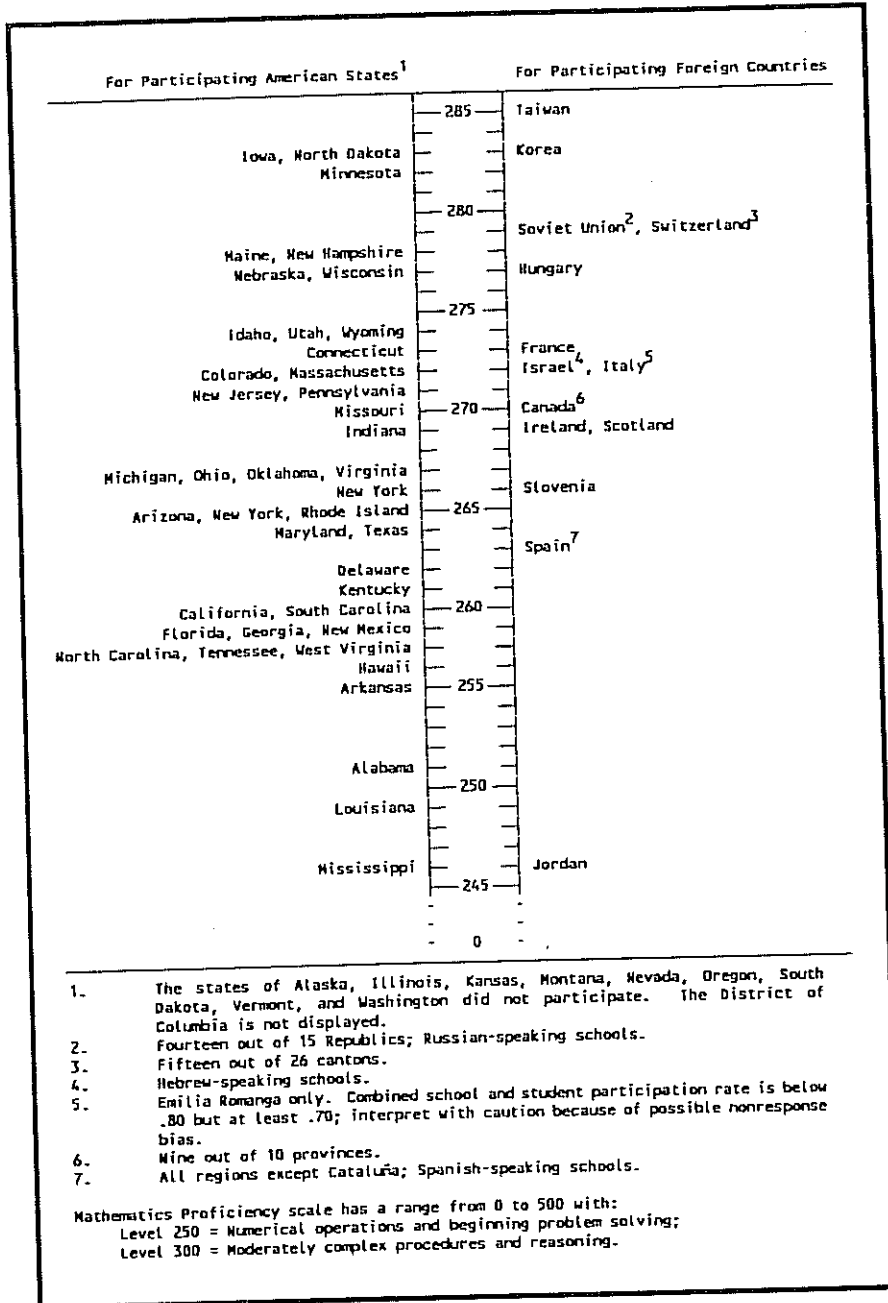
We do know of one analysis, however, that generated disaggregated data comparing the average mathematics achievements of eighth-graders from public schools in various American states with age-equivalent students from other countries. (Such data have value because poverty and annual per-student support for schools also vary among the states.) This analysis was provided by the National Center for Education Statistics and drew from a recent comparative study.⁵³ As Exhibit 2.16 shows, the average scores of eighth-graders in public schools in some high-achieving states (Iowa, North Dakota, and Minnesota) are actually just about the same today as those of the highest-achieving foreign countries included in the study (Taiwan and Korea). In contrast, poorly achieving American states (Louisiana and Mississippi) stand just about with Jordan—a struggling, not yet developed nation.

Thus, the performance of American public schools ranges from the truly remarkable to the unforgivably miserable. With a range of this magnitude, it seems fruitless to make proclamations about the “average” achievement of American schools. Instead, aggregate statistics hide more than they reveal about student achievement in our country. The real task facing America is to find ways to improve the education and lives of America’s poorest and most neglected citizens.

The Unreported News. Finally, we note with sadness that when comparative research appears in which American education looks good, somehow the press seems not to notice. As far as we know, the press paid no attention to the studies of Ian Westbury and the National Center for Education Statistics summarized in Exhibits 2.15 and 2.16. Similarly, in one of the few international comparisons ever done at the college level, John Cogan, Judith Torney-Purta, and Douglas Anderson recently found that American students know less than Japanese students about global issues when both groups enter college but that by graduation that gap has disappeared.⁵⁴ (American students tend to grow in knowledge of international affairs between their freshman and senior years, but Japanese students do not.) The press seems also to have ignored these findings.

In 1992 the IEA released a report of findings from a comparative study of reading achievement involving two hundred thousand students in thirty-

Exhibit 2.16 Average Mathematics Proficiency Scores for Public School 8th-Graders (In America) and 13-Year-Olds (In Other Countries): 1991 or 1992



—Source: *Education in the States and Nations* (National Center for Education Statistics, 1993a, pp. 56–57).

one nations. world—whi average and IEA put out paper, radio the announ only after *Et* up the story, appeared in

In a rece compared tv and Americ: to curricula predicted a students. In in mathema skills amon conceptual : formation a ageable step

If those s confirms th than are per Education 8 graduated f States and o and Develo: the Americ: enormously tries. More of their twe (Japan) is : (with 13 pe same prop nations that But have th we can tell,

In the s productivit it had been are our int stories, the cated,” “ill

one nations.⁵⁵ In this study, American nine-year-olds placed second in the world—while our fourteen-year-olds finished ninth, which was well above average and only a few points off the top. According to Gerald Bracey, the IEA put out a major press release concerning the study—but not one newspaper, radio station, or television outlet in the United States chose to cover the announcement of this study and its results. Americans learned about it only after *European* newspapers reported it and some wire services had picked up the story of high American students' achievement in reading that had appeared in the European press.⁵⁶

In a recent study, Richard Mayer, Hidetsugu Tajika, and Caryn Stanley compared two types of mathematical knowledge between samples of Japanese and American fifth-graders.⁵⁷ By comparison, Japanese students are exposed to curricula that stress more *computational* mathematics, so the researchers predicted and found greater computational achievements among Japanese students. In contrast, American fifth-grade classes stress more *problem solving* in mathematics, so the authors predicted and found greater problem-solving skills among American students. (The Americans excelled at constructing conceptual models for mathematical problems, selecting and combining information about those problems, and breaking the problems down into manageable steps.) But this study also seemed to hold no interest for the press.

If those stories do not make the case, how about the following? Evidence confirms that Americans are *far* more likely to complete higher education than are people in other Western countries. In 1993 the National Center for Education Statistics released data about the percentages of people who have graduated from colleges and universities by age twenty-two in the United States and other countries from the Organisation for Economic Co-operation and Development (OECD).⁵⁸ These data also provided comparisons among the American states (see Exhibit 2.17). Rates of college graduation also vary enormously among the states, but look at the rates for other Western countries. More than half of all American states are graduating 27 percent or more of their twenty-two-year-olds, but the *highest* percentage for a foreign country (Japan) is 26 percent. Even the lowest-ranking American states—Nevada (with 13 percent) and Alaska (with 10 percent)—are graduating roughly the same proportion of twenty-two-year-olds as are major Western European nations that are so often touted as models for American education to emulate. But have these data provoked noisy debate? They have not; indeed, as far as we can tell, media silence concerning them has been complete.

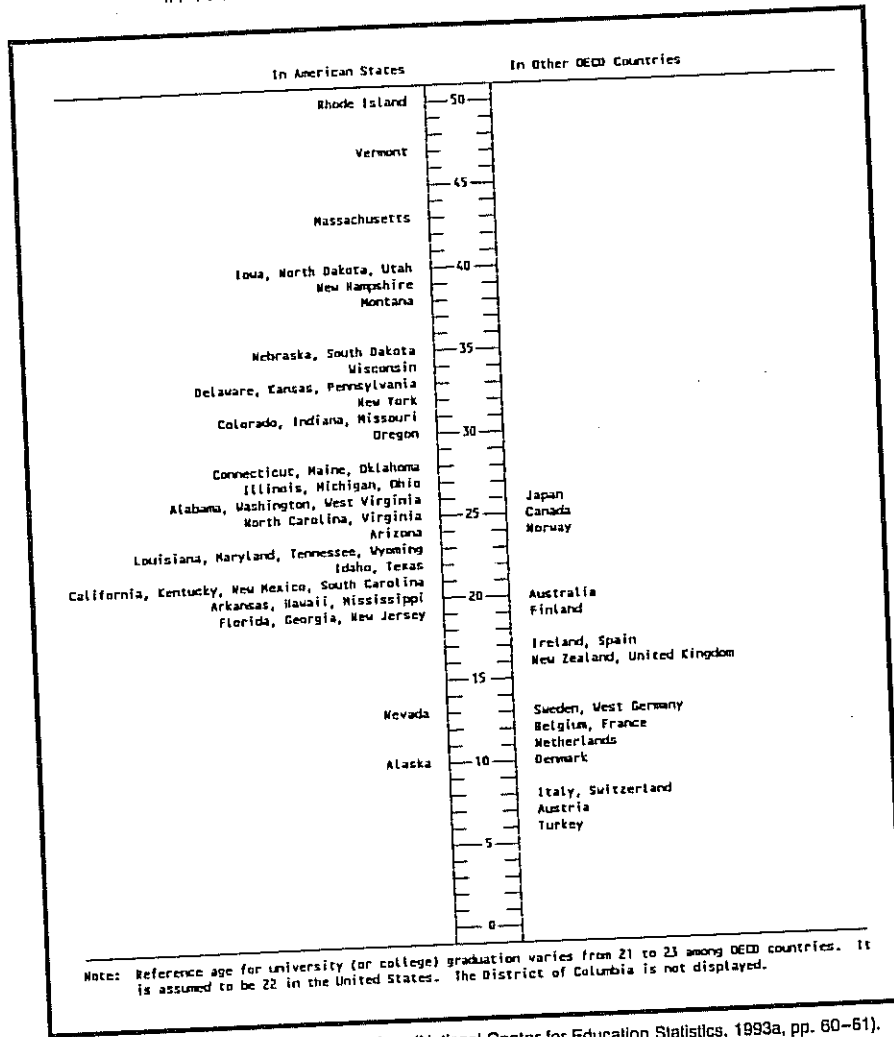
In the spring of 1993, the press *did* carry the good news that worker productivity in the United States had just hit a twenty-year high and was, as it had been for years, greatly ahead of worker productivity in countries that are our international rivals. But without questioning the incongruity of its stories, the same press thereafter continued to report on America's "uneducated," "ill-prepared," and "poorly motivated" workforce. Thus, while the

School
Coun-

ries

outh
ofelow
onse

Exhibit 2.17 Percentage of Persons Age 22 in American States and Other OECD Countries Who Were University and College Graduates in 1988



—Source: *Education in the States and Nations* (National Center for Education Statistics, 1993a, pp. 60–61).

media seem quite willing to report bad news about our schools, they are much less willing to report news about the strengths of American education.⁵⁹ One wonders, Why? Is it possible that bad news about education is thought to be more newsworthy? Or have reporters by now been so brainwashed by the critics that they cannot believe good news? Are too many of the reporters' employers close friends of powerful people who wish the public schools no

good? school

Su with : eviden fully a well. some

- I
- s
- M
- t
- M
- c
- s
- I
- h
- r
- T
- t

Th the p eviden streng Amer schoo probl Amer chapt comp

Fr Ar chi eff for

good? We don't know the answer to these questions, but America's public schools have certainly been getting a bad reputation from the media.

Summary. The myth that American schools fail badly by comparison with schools in other industrialized countries is also not supported by the evidence. Instead, when we analyze that evidence responsibly and think carefully about its implications, we discover that American schools stack up very well. And although the United States looks bad when aggregate scores from some comparative studies are examined, we should understand the following:

- Few of those studies have yet focused on the unique values and strengths of American education.
- Many of the studies' results have obviously been affected by sampling biases and inconsistent methods for gathering data.
- Many, perhaps most, of the studies' results were generated by differences in curricula—in opportunities to learn—in the countries studied.
- Aggregate results for American schools are misleading because of the huge range of school quality in this country—ranging from marvelous to terrible.
- The press has managed to ignore most comparative studies in which the United States has done well.

Thus, despite what you may have been led to believe by the critics and the press, and even though the comparative evidence to date is weak, the evidence that does exist suggests that American education has important strengths which do *not* appear in other countries. This does not mean that American education is without flaws. Indeed, achievement levels of American schools vary greatly, and that variation is a product of serious and growing problems in our society and of unconscionable differences in the funding of American schools. But, as with the other myths we have reviewed in this chapter, the myth that American education fails generally in international comparisons is balderdash.

Recapitulation

From one source after another, for the past decade we've been receiving evidence that American education is doing a mediocre job, one that ill-serves this country and our children. What is most alarming is that after a sustained period of valiant reform effort—and no small investment of resources—in the 1980s, we have so little to show for our labors and our money.

—Chester E. Finn, Jr. and Theodor Rebarber (*Education Reform in the '90s*, 1992, p. xi)

Since the early 1980s, Americans have been subjected to a massive campaign of criticism directed at their public schools and colleges. We have been told that student achievement in those institutions has slipped badly, that our achievement now lags behind that of students in other industrialized countries, and that these judgments are confirmed by numerous studies. As a result, the critics charge, American students are now being shortchanged and the nation is "at risk." Unfortunately, these charges have often been made by the White House and other prestigious sources, and they have been picked up and endlessly elaborated on in the media.

These charges are errant nonsense. If we go by the evidence, despite greatly expanded student enrollment, the average American high school and college student is now doing as well as, or perhaps slightly better than, that student did in previous years. Indeed, not only is student achievement remaining steady or rising slowly across the land, but so also is student intelligence. And when comparative-study evidence is examined carefully, that evidence also confirms impressive strengths of American education.

Unfortunately, the press, the public, and often educators themselves have not understood the evidentiary shortcomings of this massive, critical attack against American public education. So our message will come as a surprise to many people. This does not mean that all American schools do well or that the educational system in this country does not have shortcomings. Indeed, America suffers serious problems that threaten our schools, and some of those schools are miserably funded and do poorly indeed. But on average, American schools and colleges do a lot better than the critics have been claiming. As we shall see in later chapters, where those schools do fail, they usually fail because of problems imposed on them—problems that the critics have blissfully ignored.

The picture is
dismal for
commitment
schools
just achievement
system
by these
educational
such a level
educational
have achieved
Most
to debate
into the

1. Most
to
we
2. Most
we
As
3. Most
th
di

In many
public schools
a host of
we will
and a lot
This
schools
spend more
and schools