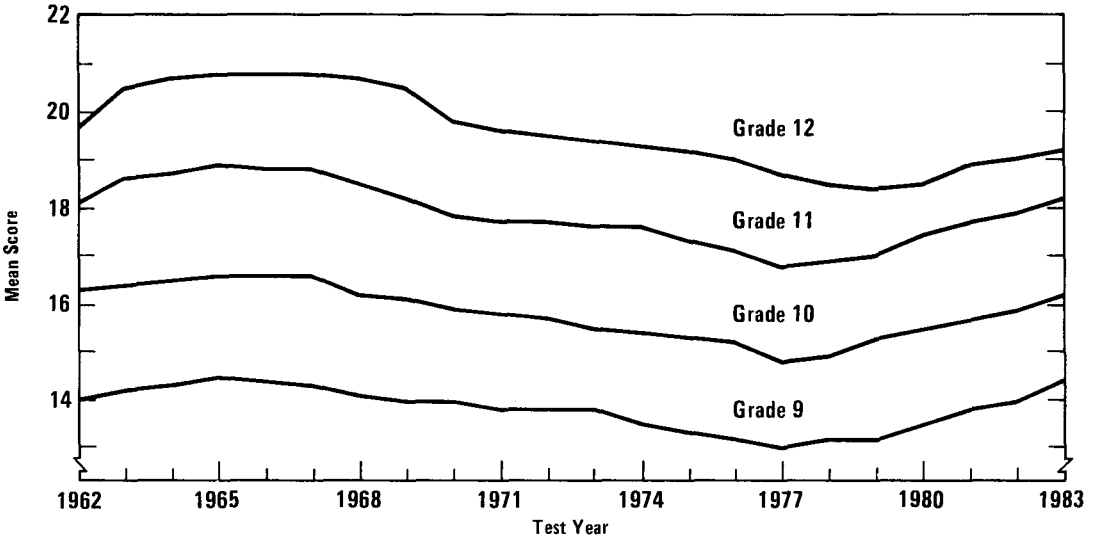
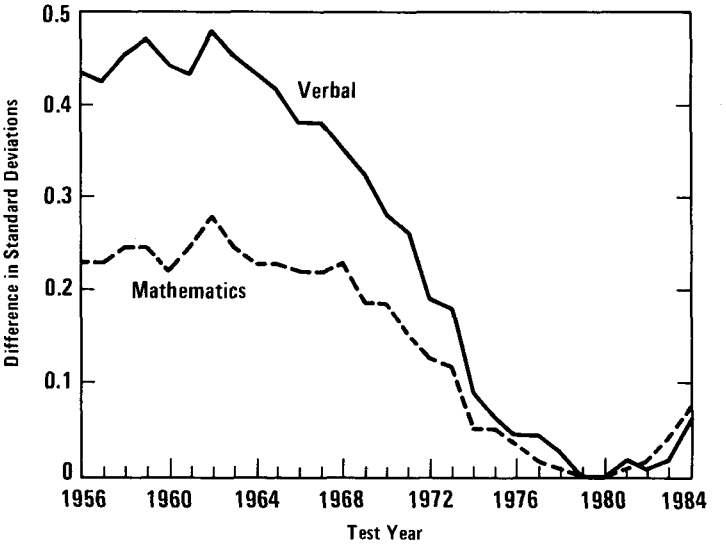
Figure III-3.
ITED Composite Scores, Iowa Only, by Test Year and Grade at Testing



SOURCE: "Mean ITED Test Scores by Grade and Subtest for the State of Iowa: 1962 to Present" (Iowa Testing Programs, undated and unpublished tabulations).

Figure III-4.
Average SAT Scores, by Subject, Differences from Lowest Year



SOURCES: CBO calculations based on Hunter M. Breland, *The SAT Score Decline: A Summary of Related Research* (New York: The College Board, 1976), Table 1; and the College Entrance Examination Board, *National College-Bound Seniors, 1985* (New York: The College Board, 1985).

(The extent to which a number of tests conform to this pattern is explained in Appendix B.)

The widespread misconception that the achievement decline ended only within the past few years thus probably stems from the greater attention paid to tests administered to high-school juniors and seniors--in particular, the SAT. The tests that showed an end of the decline taking place a decade or more ago are those given to young children, and they have been the focus of considerably less attention.

How Large Was the Decline?

The severity of the decline in achievement can be illustrated in two ways: by examining the actual level of achievement shown by typical students in each of two years (a criterion-based or absolute standard), or by comparing the achievement of a typical student in one year to the distribution of achievement in some other year (a normative standard). 12/ This section applies both standards.

The Size of the Decline Relative to a Normative Standard. The few test series reporting trend data in normative terms suggest that, at grades 6 and above, the decline averaged about 0.3 standard deviation over the entire period of the decline (see Table III-2). 13/ This average indicates that the median student at the end of the decline would have scored at about the 38th percentile at the beginning of the decline. The severity of the decline varies so greatly, however, that a single average has little value. At one extreme, the largest decline in the measures considered here was 0.55 standard deviation, placing the median student at the end of the decline roughly at the 29th percentile before the decline began. (The two largest declines, however, were on college admissions tests--the SAT and ACT--and were substantially exacerbated by changes in the composition of

12. Most often, the "typical" score is the mean or median in each year. Since the characteristics of the groups taking most tests change over time, trends in these typical scores in part reflect changes in student characteristics, rather than only changes in the achievement of a student with any given characteristics.

13. See Chapter 2 for an explanation of standard deviations.

The numbers here do not adjust the SAT trends for "scale drift," a gradual drop in the level of the difficulty of the test that led to an understatement of the SAT decline until the early 1970s. That adjustment was not made because of a lack of information about the severity or direction of any changes in difficulty since that time. If the adjustment is made, however, the conclusions of this section are unaltered; the average decline remains about 0.3 standard deviations, and the maximum decline is in the range of 0.57 to 0.60 standard deviations rather than 0.55.

TABLE III-2. SIZE OF THE ACHIEVEMENT DECLINE, INDICATED BY SELECTED TESTS AT GRADE 6 AND ABOVE
(Including only tests spanning all or nearly all of the decline)

Test	Subject	Total Decline (Standard Deviations)
Specific Tests		
SAT a/		
Largest	Verbal	0.48
Smallest	Mathematics	0.28
Iowa Grade 12 (ITED)		
Largest	Reading b/	0.40
Smallest	Mathematics	0.27
Iowa Grade 10 (ITED)		
Largest	Reading b/	0.32
Smallest	Natural Science	0.25
Iowa Grade 8 (ITBS)		
Largest	Mathematics	0.47
Smallest	Vocabulary	0.26
Iowa Grade 6 (ITBS)		
Largest	Mathematics	0.38
Smallest	Vocabulary	0.10

(Continued)

TABLE III-2. (Continued)

Test	Subject	Total Decline (Standard Deviations)
ACT		
Largest	Social Studies	0.55
Smallest	Science	-0.06 c/
All Tests in Table		
Average		0.31
Minimum		-0.06
Maximum		0.55

SOURCES: CBO calculations based on Hunter M. Breland, *The SAT Score Decline; College Board, National College-Bound Seniors, 1978* and *1985;* Iowa Testing Programs, "Mean ITED Scores by Grade and Subtest for the State of Iowa: 1962-Present" and "Iowa Basic Skills Testing Program, Achievement Trends in Iowa: 1955-1985" (unpublished and undated); Robert Forsyth, personal communication, August, 1984; A. N. Hieronymus, E. F. Lindquist, and H. D. Hoover, *Iowa Tests of Basic Skills: Manual for School Administrators* (Chicago: Riverside, 1982); L. A. Munday, *Declining Admissions Test Scores;* and American College Testing Program, *National Trend Data for Students Who Take the ACT Assessment.*

NOTE: Alternate grades (7, 9, 11) omitted for clarity.

a. SAT scores are not adjusted for scale drift. Research indicates that the first part of the decline is understated by perhaps 0.09 standard deviations because of scale drift. The extent and direction of scale drift over the past decade is not yet known, however.

b. This reflects the "Interpretation of Literary Materials" test. Reading skills also are tapped by the other tests in the ITED battery.

c. Negative numbers represent an increase in average scores.

the test-taking group.) 14/ At the other extreme, the ACT natural science test actually showed a trivial increase during the years of the general decline. Thus, a different mix of tests--or a larger and more representative sample of tests--might have yielded a very different average size of the decline.

Some of the variability in the size of the decline stems from known causes, such as the age of the students tested and changes in the composition of the groups of students taking the test. On the other hand, much of the variation appears to stem from unknown factors or from considerations that lie largely outside of the scope of this report, such as decisions about the specific skills and knowledge to be tested.

The Size of the Decline Relative to an Absolute Standard. Although the apparent severity of the decline varies with the absolute achievement criterion chosen, the average decline was clearly large enough by many standards to be educationally significant.

The best criterion-based gauge of the achievement decline is probably the National Assessment of Educational Progress (NAEP). The NAEP reflects representative samples of the national population of students, tests students at the elementary, junior-high, and senior-high levels, and encompasses a wide array of substantive areas and types of skills. Moreover, actual test items from all of the NAEP assessments have been published, along with the percentages of students of different ages answering each item correctly. This information provides an intuitively clear view of students' levels of achievement. 15/

Even the NAEP, however, should be used to illustrate the types of skills that deteriorated rather than to indicate the total magnitude of the decline. Because of the timing of NAEP assessments, most of them understate the severity of the decline, in some instances probably by a very large margin. The NAEP began with a science assessment in 1969, with initial assessments in other subjects starting over the following several

14. Advisory Panel on the Scholastic Aptitude Test Score Decline, *On Further Examination* (New York: The College Entrance Examination Board, 1977); L. A. Munday, *Declining Admissions Test Scores*. The impact of compositional changes is discussed in Congressional Budget Office, *Educational Achievement: Explanations and Implications of Recent Trends* (forthcoming).

15. While annually equated tests provide much clearer information on trends, no such tests have been tabulated in a way that facilitates comparison with an absolute achievement criterion.

years. The most recent published assessments were largely carried out between 1976 and 1981. Therefore, the NAEP trends exclude varying portions of the early part of the decline and probably often mask the later decline by mixing with it upturns in achievement that have occurred in recent years. 16/

Although the NAEP tests students at ages 9, 13, and 17, this section describes the results among 17-year-olds, because the trends among 17-year-olds are likely to include fewer, if any, years of the recent upturn in achievement. 17/ (Comparable information on ages 9 and 13 appears in Table III-3.)

Mathematics. Between 1972 and 1977, the proportion of NAEP mathematics items answered correctly by 17-year-olds dropped from 64.0 to 60.4 percent (see Table III-3). While this decline appears modest, it occurred over a time span that was probably less than half of the total period of decline and also masks more substantial deterioration of performance on certain important types of items. 18/ In addition, the rate of success on certain types of items was remarkably poor in both years. One NAEP computation item, for example, asked: "Express 9/100 as a percent." The proportion of 17-year-olds answering this item correctly dropped eight percentage points over the five years, from 61 percent to 53 percent. Similar results were obtained by a problem that asked: "A hockey team won 5 of the 20 games it played. What percent of the games did it win?" Another problem required students to use a simplified electrical bill to determine the cost per kilowatt if 606 kilowatts produced a bill of $9.09. The proportion of students succeeding on this item fell from 12 percent in 1973 to 5 percent in 1978. 19/

---

16. Because NAEP assessments are carried out at intervals of four or five years, the ends of the decline in each of them cannot be pinpointed. This precludes estimating any recent increase in each series and disentangling it from the estimates of the preceding downturn.

17. In interpreting the examples given below, it is important to bear in mind that only 17-year-olds still in school were tested in the National Assessment. As a result, the NAEP results are likely to overestimate--perhaps by a sizable margin--the average level of achievement attained by the entire cohort of 17-year-olds.

18. The subsequent interval from 1977 to 1981 showed little change, but it probably brackets the end of the decline and therefore includes some of the subsequent upturn.

19. These and the following mathematics examples are taken from National Assessment of Educational Progress, *Changes in Mathematical Achievement, 1973-1978* (Denver: NAEP/ Education Commission of the States, 1979).

TABLE III-3. SUMMARY OF NATIONAL ASSESSMENT RESULTS IN THREE SUBJECTS, AGES 9, 13, AND 17
(Average percent of items correctly answered)

Subject	Age 9	Age 13	Age 17
Mathematics a/			
1972	56.7	58.6	64.0
1977	55.4 b/	56.6 c/	60.4 c/
1981	56.4	60.5 c/	60.2
Reading d/			
1970	64.0	60.0	68.9
1974	65.2 c/	59.9	69.0
1979	67.9 c/	60.8	68.2
Science			
1969	61.0	60.2	45.2
1972 e/	59.8 c/	58.5 c/	42.5 c/
1972 f/	52.3	54.5	48.4
1976	52.2	53.8	46.5 c/

SOURCES: CBO calculation based on National Assessment of Educational Progress, *Three National Assessments of Reading* (1981), Tables 2, 4, and 6. Mathematics: *The Third National Mathematics Assessment: Results, Trends, and Issues* (1983), Tables 5.1 and 5.2, and *Mathematical Technical Report: Summary Volume* (1980), Tables 2, 3, and 4; and *Three National Assessments of Science* (1978), Table A-1 (Denver: NAEP/Education Commission of the States).

a. 1977 and 1981 scores reflect all items used in those two assessments. 1972 scores are obtained by subtracting from 1977 scores the change between 1972 and 1977 on all items used in those two years.

b. Change from preceding test marginally significant, p less than .10.

c. Change from preceding test statistically significant, p less than .05.

d. All scores reflect all items used in all three years.

e. Reflects only test items shared with 1969.

f. Reflects only test items shared with 1976.

Student achievement also dropped on certain NAEP items that were less tied to concrete applications. For example, the proportion of 17-year-old students correctly finding a missing numerator in an equivalent fraction fell from 82 percent to 72 percent. The proportion who could solve for x and y in a system of linear equations dropped by a third, from 18 percent to 12 percent.

On the other hand, success rates on some items did not decline--an optimistic note that is tempered by the fact that in many instances the rate was poor in both years. For example, in both 1973 and 1978, about 20 percent of students successfully graphed the equation $y = 2x + 1$. About 15 percent and 12 percent could identify the slope and intercept, respectively, of the equation $2y = 5x - 8$. Five percent ascertained the equation of a line when both the x- and y-intercepts were given.

Reading. In contrast to mathematics, the first three NAEP reading assessments showed no substantial overall decline in the achievement of 17-year-olds (see Table III-3). This pattern is inconsistent with a variety of other tests that showed substantial declines in reading and reading-related skills. The results of those other tests, however, have not been published in a form that permits comparison with a concrete achievement criterion.

On the other hand, a decline was apparent in one of the specific reading skill areas tapped by the NAEP--inferential comprehension (that is, comprehension that requires going beyond the information explicitly stated in the question). This discrepancy is discussed in a later section.

Science. Over the seven-year span covered by the first three NAEP assessments of science--1969, 1972, and 1976--the average score of 17-year-olds dropped 4.6 percentage points, or about 10 percent (see Table III-3).

As in the case of the mathematics assessment, the low success rate on certain items is as striking as the decline. One NAEP item, for example, asked, "Which of the following happens when any combustion reaction takes place?" The correct choice--that heat is evolved--was selected by about 68 percent of 17-year-olds in 1969 and by about 54 percent in 1977. Another item asked for explanation of the statement that the relative humidity is 50 percent. About 47 percent of students in 1969 and 42 percent in 1977

selected the correct answer--"The atmosphere contains half as much water as it could at its present temperature." 20/

Social Studies. The NAEP citizenship and social studies assessments in 1968, 1971, and 1975 showed sizable declines in the proportion of 17-year-olds correctly answering some items assessing knowledge of the Constitution, the structure and function of government, the political process, and international affairs. A smaller number of items, however, showed increases.

In one example, the proportion of students answering that a statement of civil rights can be found in the Constitution dropped from 85 percent to 81 percent between 1971 and 1975. 21/ The proportion correctly answering the question "The Congress of the United States is made up of two parts. One part is the House of Representatives. What is the other part?" fell from 94 percent to 88 percent from 1968 to 1975. (The proportion choosing the most popular incorrect answer--the Supreme Court--doubled to 8 percent during that period.) The proportion recognizing that the Congress was part of the legislative branch of government dropped during the same time, from 84 percent to 74 percent. Fifty-four percent of 17-year-olds in 1968, but only 35 percent in 1975, recognized that the circumstance of a state having more Senators than Representatives occurs as a result of low population. The proportion able to define "democracy" declined from 86 percent to 74 percent between 1968 and 1975.

THE RECENT UPTURN IN ACHIEVEMENT

Since the end of the achievement decline, the general trend has been a marked upturn in average achievement. In some instances, the rate of increase has been comparable to or even greater than the rate of decrease during the later years of the decline, and average scores on some tests have approached or exceeded their predecline high points. Moreover, the pattern

20. National Assessment of Educational Progress, *Three National Assessments of Science: Changes in Achievement, 1969-77* (Denver: NAEP/Education Commission of the States, 1978).

21. This and the following examples are taken from National Assessment of Educational Progress, *Changes in Political Knowledge and Attitudes, 1969-76* (Denver: NAEP/Education Commission of the States, 1978).

of the trends among tests administered at different ages suggests that some of the test batteries that have seen only a modest upturn to date--most notably, the SAT--might show marked increases in the next several years.

In contrast, there are a number of important exceptions to this optimistic picture. Scores on the American College Testing Program college admissions tests have yet to turn up substantially. A statewide assessment program in Pennsylvania has shown stable scores in the lower grades and slight deterioration at the secondary level in recent years.22/ The California statewide assessment also has shown no upturn among seniors, though it has shown increases in the lower grades. 23/

Much of the variation in recent trends appears linked to the age of the students: tests given to older students have generally increased less in total than have those administered to younger children. At one extreme, some tests administered in the elementary grades have risen to their highest levels on record--a span of as much as three decades. At the other pole, the generally better known tests administered in the high school grades (such as the SAT) have generally shown more modest gains.

The smaller total upturn to date in the higher grades appears to reflect the shorter time since the upturn began in those grades, rather than a lesser rate of improvement. The upturn, like the end of the decline, shows a cohort pattern, and fewer of the cohorts producing rising scores have yet reached the higher grades. (The relationships between age and the subsequent upturn are discussed further in Chapter IV.)

This pattern suggests that scores on tests administered in the higher grades might rise further in the coming years. That is, the cohorts responsible for the most recent rise in scores in the lower grades might be expected to produce similar gains as they move through the higher grades. The cohort pattern notwithstanding, however, any number of factors could cause future trends in the higher grades to diverge from the recent trends produced by those same cohorts in the earlier grades.

---

22. Robert Coldiron, Pennsylvania State Department of Education, personal communication, January 1985.

23. California State Assessment Program, unpublished tabulations.

Has the Upturn Ended?

The most recent National Assessment of reading found that the average reading proficiency of nine-year-olds was largely unchanged between 1979 and 1983, while the achievement of older students continued to rise.[24] That is, the birth cohort of 1974 showed no gain over the birth cohort of 1970.

Given the cohort pattern evident in the achievement upturn, this pattern--if it appears on other tests and is maintained--suggests that the upturn is, for the moment, over in the youngest age groups and that it will end fairly soon in the higher grades (as the birth cohorts that were nine years old between 1979 and 1983 pass through the grades). Tests administered to eighth graders would be expected to level off in the 1983 to 1987 period, while scores of seniors would level off between 1987 and 1991.

Whether this leveling off is a general phenomenon, however, is unclear. No other national data are available to test it, and state-level data are inconsistent. The proportion of New York third-grade students passing the state reference points in mathematics and reading, for example, has been stable since the 1970 and 1971 birth cohorts (see Figure B-5 in Appendix B). On the other hand, average scores in the elementary grades in Iowa have continued to rise, even in the most recent (1984-1985) year of data (see Figure III-2). In the next several years, National Assessments will take place in other subject areas, which will provide nationally representative data indicating whether this leveling off is a general occurrence.

DIFFERENCES IN TRENDS AMONG TESTS

Recent achievement trends have varied greatly from one test to another. For example, comparisons of recent trends on the SAT, the ACT, and standardized tests given to high school juniors and seniors as a whole show many discrepancies from one test to another (see Table III-4). This variation indicates that no single test, taken alone, is an adequate indicator of overall achievement trends. Indeed, in the absence of a clear understanding of the variations in the trends from one test to another, even a few tests taken together cannot always be assumed to be a sufficient indicator.

---

24. National Assessment of Educational Progress, *The Reading Report Card: Progress Toward Excellence in Our Schools* (Princeton: NAEP/Educational Testing Service, 1985).

This variation apparently reflects differences both among the tests themselves and among the students taking them. The precise role of each is unclear, however, and some of the specific differences between tests are hard to explain. For example, although the ACT sample is in some important respects comparable to the SAT group and underwent some of the same compositional changes as affected the SAT, the trends on the two tests are markedly different. 25/ Conversely, although the Iowa ITED is substantively similar to the ACT and was presumably free of many of the compositional changes that biased the SAT and ACT trends, it showed total declines roughly as large as those shown by the SAT (see Table III-2).

Subject Areas

The relative severity of the decline in different subject areas has been the focus of considerable discussion, in terms of both explanations of the trends and debates about appropriate responses. Debate has focused not only on specific subject areas, but also on two broad categories of subjects: those primarily taught "directly" in school, and those that are to a substantial degree taught "indirectly," both in school and elsewhere. 26/ Some people would argue, for example, that certain mathematical skills--such as converting fractions to decimals or solving algebraic equations--are taught primarily in school through formal instruction and drill. In contrast, a larger proportion of vocabulary knowledge is presumably learned as an incidental result of daily experience at home and elsewhere. For this reason, a larger decline in the "indirectly" taught subjects might imply that the decline was attributable more to changes in student characteristics or to broad social changes than to changes in schooling, while larger declines in "directly" taught subjects would implicate schooling. 27/

---

25. Compositional changes affecting ACT means are discussed in Munday, *Declining Admissions Test Scores*.

26. Donald Rock and others, *Factors Associated with Decline of Test Scores*, p. 6.

27. While few people would argue with the idea that students learn a larger proportion of their vocabulary than of their mathematical skills outside of school, the observed relationships between achievement in different subject areas and home and school characteristics are not clear-cut. For example, an analysis of the relative size of home and school effects on achievement in several countries found that schooling effects were indeed larger in science than in reading among 10-year-olds but not among 14-year-olds (James S. Coleman, "Methods and Results in the IEA Studies of Effects of School on Learning," *Review of Educational Research*, vol. 45, Summer 1975, pp. 355-386, Tables 2 and 3.)

TABLE III-4. RECENT TRENDS ON STANDARDIZED TESTS AMONG HIGH SCHOOL SENIORS AND JUNIORS, WITH TRENDS OVER THE SAME PERIODS ON THE SAT a/

Test	Subject	Change (Standard Deviations)
	1970 to 1983	
National Assessment	Reading	.10
SAT	Verbal	-.26
ACT	English	.02
ITED-Iowa Grade 12	Vocabulary b/	-.08
	Reading c/	-.12
SAT	Mathematics	-.14
ACT	Mathematics	-.23
ITED-Iowa Grade 12	Mathematics b/	-.03
	1971 to 1979	
NLS to HSB	Vocabulary	-.22
	Reading	-.21
SAT	Verbal	-.26
NLS to HSB	Mathematics	-.14
SAT	Mathematics	-.15
	1970 to 1981	
Illinois Decade Study d/	English 1	-.38
	English 2	-.49
SAT	Verbal	-.25

(Continued)

TABLE III-4. (Continued)

Test	Subject	Change (Standard Deviations)
	1970 to 1981 (cont'd.)	
Illinois Decade Study	Mathematics 1	-.05
	Mathematics 2	-.26
SAT	Mathematics	-.14

SOURCES: CBO calculations based on National Assessment of Educational Progress, *The Reading Report Card;* Albert Beaton, NAEP/Educational Testing Service, personal communication, December 1985; Hunter M. Breland, *The SAT Score Decline;* Table 1, College Board, *National College-Bound Seniors, 1978 and 1985;* L. A. Munday, *Declining Admissions Test Scores;* and American College Testing Program, *National Trend Data for Students Who Take the ACT Assessment;* Iowa Testing Programs, "Mean ITED Scores by Grade and Subtest for the State of Iowa: 1962-Present;" Robert Forsyth, Iowa Testing Programs, personal communication, August, 1984; Donald A. Rock, Ruth B. Ekstrom, Margaret E. Goertz, Thomas L. Hilton, and Judith Pollack, *Factors Associated with Decline of Test Scores of High School Seniors, 1972 to 1980* (Washington, D.C.: Center for Statistics, U.S. Department of Education, 1985); *Student Achievement in Illinois, 1970 and 1981* (Springfield: Illinois State Board of Education, 1983).

a. The dates used in each set reflect the longest portion of the 1970-1983 period for which data are available. The NLS/HSB and Illinois Decade data are available only for the periods indicated. Comparisons extending past 1979 generally include a period of increasing scores.

b. These small changes in the ITED reflect substantial declines that were nearly offset by gains since 1978 and 1979.

c. This reflects the "Interpretation of Literary Materials" test. Reading is also tested on other tests in the ITED battery.

d. High school juniors only. SAT comparisons are therefore one year later.

Among the tests assessed here, no single subject area consistently showed the largest drop, and the decline was not consistently larger among either directly or indirectly taught subjects. In a majority of the tests, the drop was largest on language-related tests such as verbal reasoning, language usage, vocabulary, and reading. The exceptions were frequent enough, however, to suggest that this pattern is more a reflection of the particular tests than an underlying characteristic of the achievement decline.[28] Indeed, a different assortment of tests--if more were available--might show a very different aggregate ranking of the decline in different subject areas.

Thus, for example, language-related tests showed the largest drops on the SAT, a nationally representative comparison of high school seniors in 1971 and 1979 (the NLS and HSB comparison), and in some of the Iowa data. Conversely, mathematics showed the steepest decline in other Iowa data and in the national normings of the California Achievement Test. Moreover, some of the language-related tests that showed particularly large declines (such as the vocabulary test in the NLS and HSB comparison) tap indirectly taught subjects, while others (such as the language test in the ITBS data and the ITED expression test) are clearly much more reliant on formal instruction. (For more detail on the relative size of the decline in different subject areas, see Appendix C.)

Underlying this seeming lack of consistency is the fact that achievement in any one subject can be defined--and measured--in many different ways, and the variations in measurement can be large enough to create very different trends. Thus, to speak of "the decline in mathematics achievement" is misleading. It is more accurate to speak of the decline in the mathematics skills measured by a specific test, and one should bear in mind that other tests might yield very different trends.

Trends in average mathematics achievement of Iowa students clearly illustrate the effect of test differences on the severity of the decline.[29]

---

28. This discussion reflects only tests for which standard deviations are available, since the trends in different subject areas are made comparable by expressing them as fractions of a standard deviation. The National Assessment is therefore excluded, since standard deviations from previous assessments were not all retained by the NAEP staff. (Lawrence Rudner, Office of Educational Research and Improvement, U.S. Department of Education, personal communication, January 1985).

29. Since most students in Iowa are tested with the ITBS through grade 8 and with the ITED in grades 9 through 12, differences between trends in Iowa on the grade 8 ITBS and the grade 9 ITED reflect little other than the differences in the tests themselves. The scores are based on almost the same group of students at nearly the same point in their school careers.

Over the entire period of the decline, the eighth-grade Iowa ITBS dropped substantially more in mathematics than in other subjects. In contrast, the ninth-grade Iowa ITED showed somewhat less decline in mathematics than in social studies or reading (the "interpretation of literary materials"). Over the whole period, the mathematics decline on the grade-eight ITBS was nearly half a standard deviation, or about .036 standard deviations per year. On the ITED, the total mathematics decline and the annual rate of the decline were both roughly half as large (see Figure III-5). The explanation of this difference might lie in the construction of the tests; the ITBS is roughly split between concept items (which are highly curriculum bound) and applications items, while the ITED places much greater emphasis on the latter. 30/

Level and Type of Skill

Evidence of the trends in different types and levels of skills is of two types: direct comparisons of different items within individual tests, and indirect inferences from comparisons of different tests--such as those in different subject areas or given at different grade levels. Direct comparisons of items can be carried out on any test, but little such analysis is currently available. Indirect inferences are therefore also noted in this section.

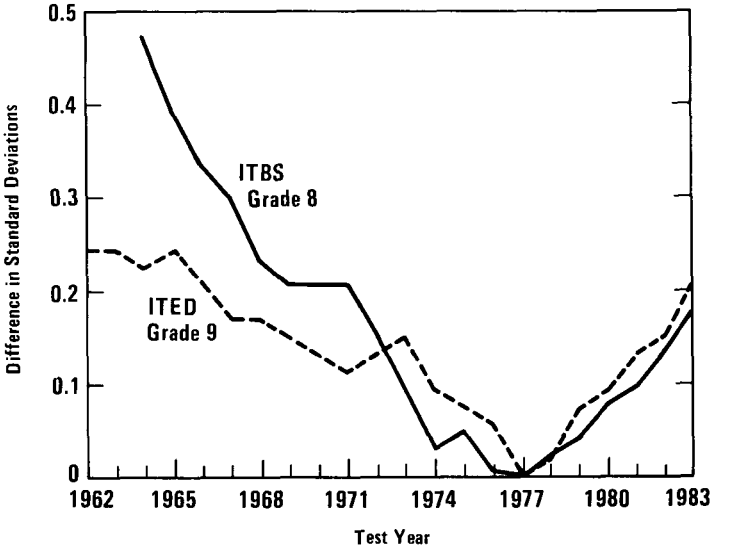
The Decline. That the overall drop in achievement entailed sizable declines in higher-level skills, such as inference and problem-solving, is beyond question. 31/ The extent to which declines occurred in more basic skills, such as simple arithmetic computation, is less clear. While some tests showed substantial declines in basic skills, other indices of basic skills showed little or no drop. In the aggregate, the evidence suggests that declines in the more basic skills might have been generally less severe than in higher-order skills, but not without exception.

---

30. Robert Forsyth, Iowa Testing Programs, personal communication, February 1985.

31. While the evidence leaves no doubt that substantial declines occurred in some higher-level skills, not all higher-level tests showed declines. The most notable exception is the Project TALENT 15-year retests, which showed increases in abstract reasoning and creativity in grades 9 through 11 between 1960 and 1975 (Table C-1 in Appendix C). This exception, however, might be artifactual. The starting point of the comparison --1960--antedated the predecline peak in achievement, thus confounding earlier growth in achievement with the decline (Cleary and McCandless, *Summary of Score Changes (in Other Tests)*). In addition, the 15-year retest suffers from two serious threats to validity and representativeness: a very small sample (only 17 schools), and meager assessment of changes in school characteristics that might bias the results.

Figure III-5.
Iowa Mathematics Achievement, Differences from Lowest Year



SOURCES: CBO calculations based on "Mean ITED Test Scores by Grade and Subtest for the State of Iowa" and "Iowa Basic Skills Testing Program, Achievement Trends in Iowa: 1955-1985" (Iowa Testing Programs, undated and unpublished tabulations); Robert Forsyth, Iowa Testing Programs, personal communication, April 1984; and A. N. Hieronymus, E. F. Lindquist, and H. D. Hoover, *Iowa Tests of Basic Skills: Manual For School Administrators* (Chicago: Riverside, 1982).

A greater decline in higher-order skills is apparent in the performance of 17-year-olds on the first two NAEP mathematics assessments (1972-3 and 1977-8), which span the last years of the decline. Performance on these tests was tabulated separately for four types of skills:

- Knowledge: "recall of facts and definitions," including facts of the four basic arithmetic operations and measurement.

- Skills: "the ability to use specific algorithms and manipulate mathematical symbols." This domain includes "computation with whole numbers, fractions, decimals, (and) percents...; taking measurements; converting measurement units; reading graphs and

tables; and manipulating geometric figures and algebraic expressions." 32/

- Understanding: items "implying a higher level of cognitive process than simply recalling facts or using algorithms." Items in this domain required explanation or illustration of various skills and "transformation" of mathematical knowledge.

- Applications: items requiring the use of the preceding three types of skills, usually in problem-solving. 33/

Average performance in the simplest domain--mathematical knowledge--did not change at all during the five-year interval (see Table III-5). (An increase in performance on items involving metric measures offset a relatively small decline in the rest of this domain.) Both of the two highest levels--understanding and applications--showed declines. Moreover, the average performance in the applications domain was very low in both years. 34/ The "skills" area showed a comparably large decline, but within that area, the drop tended to be largest on the more complex items. 35/

The second international mathematics assessment by the International Association for the Evaluation of Educational Achievement (IEA) yielded results in grade eight that are comparable to the NAEP in this respect. Average achievement in grade eight fell over the 18 years between the first and second assessments, but the declines were greater "for more demanding comprehension and application items than they were for computation items." 36/ On the other hand, the same assessment found precisely the

---

32. At the simple pole, the "skills" domain incorporates items that would be considered "basic skills" by all observers--for example, simple arithmetic operations. At the other pole, it subsumes some fairly complex operations, such as solving a system of linear equations for x and y and solving quadratic equations.

33. National Assessment of Educational Progress, *Changes in Mathematical Achievement, 1973-78*, p. xi.

34. *Ibid.*, pp. 12-15.

35. *Ibid.*, pp. 4-9.

36. F. Joe Crosswhite, John A. Dossey, Jane O. Swafford, Curtis C. McKnight, Thomas J. Cooney, and Kennth J. Travers, *Second International Mathematics Study: Summary Report for the United States* (Champaign: Stipes Publishing Co., 1985), p. xi. Given the timing of the two assessments and the age of the students, the eighth-grade trends in the international assessment probably combine several years of increasing achievement with a longer, previous period of decline.

TABLE III-5. NAEP MATHEMATICS CHANGES 1972-1977, AGE 17, BY AREA (Average percent of items correctly answered)

Area	1972	1977	Change
Total	52	48	-4 a/
Knowledge b/	63	63	0
Knowledge c/	63	62	-2 a/ d/
Skills	55	50	-5 a/
Understanding	62	58	-4 a/
Applications	33	29	-4 a/

SOURCE: NAEP, *Changes in Mathematical Achievement, 1973-78*, Tables 1-4, 6, and 7.

a. Statistically significant, *p* less than .05.

b. Including metric measures.

c. Excluding metric measures.

d. Components do not yield stated change because of rounding.

opposite pattern among 12th-grade students: an increase in achievement, much of which "was seen in the more demanding comprehension questions and, for the calculus students, at the even more demanding application level." 37/ The 12th-grade results, however, were in large part a reflection of the performance of calculus students, who constitute a small and select segment of the senior class and whose performance may therefore say little about that of high school students in general.

Evidence of a greater decline in higher-order skills also appeared in the NAEP reading assessments. As noted earlier, 17-year-olds showed little total change in reading between the 1970-71 and 1979-80 assessments. The small (and statistically insignificant) change in total reading performance,

37. *Ibid.*, p. xi. Whether these gains reflect favorable trends during the period of the general achievement decline, an earlier or particularly sharp upturn, or both remains unclear. Because the final test was administered only a few years after the end of the general decline among seniors, however, it suggests that progress during the years of general decline played a role.