

**CENTER FOR EDUCATION REFORM
AND
EMPOWER AMERICA**

**ACHIEVEMENT IN THE UNITED STATES:
PROGRESS SINCE A NATION AT RISK?**

By:

Pascal D. Forgione, Jr., Ph.D.
U.S. Commissioner of Education Statistics
National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, N.W., Room 400
Washington, DC 20208
202-219-1828 (Telephone)
202-219-1736 (Fax)
Pascal_Forgione@ed.gov (email)

NCES World Wide Web Home Page: <http://www.nces.ed.gov>

April 3, 1998

TABLE OF CONTENTS

Highlights.....	3
Introduction.....	5
I. Performance Over Time.....	6
Long-Term Trends in Science, Mathematics, and Reading.....	6
Subgroup Performance on NAEP.....	7
Framework-based Assessments in Mathematics, Reading, and Science.....	8
II. International Comparisons.....	12
International Comparisons of Mathematics and Science.....	12
Lessons From TIMSS.....	14
International Comparisons of Reading.....	16
International Perspective on Labor Force Proficiency.....	17
III. Changes in Students' Behavior Since <i>A Nation at Risk</i>.....	17
Dropping Out of School.....	18
Educational Aspirations and College Attendance.....	18
Coursetaking Patterns in High School.....	18
Conclusion.....	19
Figures.....	20

Highlights

Student Achievement Over Time

- Long-term trends in science and mathematics show declines in the 1970s and early 1980s, followed by modest increases. For example, the mathematics score averages of 17-year-olds declined from 1973 to 1982, then increased to a level in 1996 similar to the 1973 level.
- Long term trends in reading achievement show minimal changes across the assessment years. In 1996, the average reading score for 9-year-olds was higher than it was in 1971. Thirteen-year-olds showed moderate gains in reading achievement; in 1996, their average reading score was higher than that in 1971. There was an overall pattern of increase in reading scores for 17-year-olds, but the 1996 average score was not significantly different than in 1971.
- Many states have had increases in mathematics performance in recent years. Eighth-graders in 27 out of the 32 jurisdictions participating in both the 1990 and 1996 assessments showed an *increase* in their average scale scores.
- Despite these widespread increases in performance, large variations in state mathematics achievement persist. The proportion of eighth-graders performing at a *Basic* or above level ranged from 36 percent in Mississippi to 77 percent in Maine and North Dakota and 78 percent in Iowa.
- The mathematics and science achievement gap between white, black, and Hispanic students, has narrowed somewhat since *A Nation at Risk*. Blacks and Hispanics in each of the age groups tested (9, 13, and 17-year-olds) tended to make larger gains than whites during this period. Paradoxically, the achievement gains of each of these major subgroups are larger than that for the nation as a whole because of compositional changes in the student population. In particular, the lowest scoring subgroups represent a greater share of the population in 1996 than in earlier years.

International Comparisons

- Data from the Third International Mathematics and Science Study (TIMSS) suggests that the relative international standing of U.S. students declines as they progress through school. In both subject areas, our students perform above the international average in grade 4, close to the international average in grade 8, and considerably below it in grade 12.
- In twelfth-grade, the achievement scores of both our overall student population tested on general mathematics and science knowledge, and of our more advance students tested in mathematics and physics, were well below the international average.
- Findings from TIMSS suggest that many of the “cure-alls” recommended in the past are not associated with high performance in all nations. While strategies such as more homework, more seat time, and less television may be important in improving the

achievement of individual students and schools, they do not appear to be potent variables in explaining cross-national student achievement differences.

- U.S. students perform relatively well in reading compared with their international counterparts. Out of 27 countries in fourth-grade and 31 countries in ninth-grade, only Finland's achievement was significantly higher than that of the U.S.
- TIMSS data do encourage us to focus on rigorous content, focused curriculum, and good teaching as critical to improved national performance. For example, while most countries introduce algebra before high school, in the U.S. only 25 percent of students take algebra before high school. Similarly, fully 90 percent of all U.S. high school students stop taking mathematics before getting to calculus. And 55 percent of physical science teachers in this country (i.e., teachers of chemistry, physics, earth science or physical science) lack either a major or minor in their teaching sub-field.
- The United States has close to 20 percent of the adult population at both the high and low ends of the literacy scale. In contrast, European countries tend to have an adult population with skills concentrated in the middle literacy levels.
- Workers with higher literacy scores are unemployed less and earn more than workers with lower literacy scores.

Changes in Student Behavior Since A Nation at Risk

- The dropout rate has declined since *A Nation at Risk*, particularly for blacks. The Hispanic dropout rate remains much higher than for black or whites and has not changed significantly since 1982. However, the dropout rate for Hispanic immigrants is much higher (44 percent), than for first-generation Hispanics born in the U.S. (17 percent).
- The educational aspirations of high school seniors increased substantially between 1982 and 1992. In 1992, 69 percent of seniors said that they hoped to graduate from college, compared with 39 percent of 1982 seniors.
- There has been a marked increase in the number of mathematics and science courses taken by high school graduates. Between 1982 and 1994, the percentage of high school graduates completing the "New Basics" curriculum (4 years of English, 3 of social sciences, 3 of sciences, and 3 of mathematics) rose from 14 percent to 50 percent.

Introduction

Fifteen years ago when *A Nation at Risk* was released some critics charged that the report was long on conclusions and short on evidence. One observer argued that the report's subtext was the appalling lack of reliable, national education data at the disposal of policy analysts and policymakers at that time. Today, as I stand before you as the Commissioner of Education Statistics I can say -- in terms of data -- things have improved. And some of the people responsible for expanding the Nation's investment in education data are in this room today. As a result, we now have a much clearer picture of how well American schools and their students are faring.

To ask if today's students are as smart as students used to be – if they know more or can do more – invokes the most traditional and simplest form of benchmarking; it compares performance today by the standard of performance in the past. That is the main question I will address today – to ask if students are performing better by presenting data from the National Assessment of Educational Progress (NAEP) which looks at national and state performance over time. What we shall see is that the news is mixed.

But there are other ways to ask the general question “how are we doing?” Policymakers often ask if American students are doing as well as they should or as well as they can. International comparisons present an alternative kind of benchmark for gauging overall performance and are probably the most important indicator to business leaders. Comparisons of academic performance among our major economic partners are leading indicators for employers who must compete in a global economy. International comparisons are the second group of data I want to present here today, and for that I will draw primarily from the Third International Mathematics and Science Study (TIMSS), the International Reading Literacy Study (IRLS) and the International Adult Literacy Survey (IALS). These data also paint an uneven picture of our relative educational standing.

Finally, I will present data on how students have responded to the call for better performance and higher standards. We shall see that students have changed their behavior since *A Nation at Risk*: they are more likely to graduate from high school, have higher educational aspirations, and take more academic courses.

I. Performance Over Time

Long-Term Trends in Science, Mathematics, and Reading

Measuring students' academic performance has been the purpose of the National Assessment of Educational Progress (NAEP) since its inception in 1969. Students in both public and nonpublic schools have been assessed in various subject areas on a regular basis. In addition, NAEP collects information about relevant background variables to provide an important context for interpreting the assessment results and to document the extent to which education reform has been implemented.

NAEP enables us to monitor trends in academic achievement in core curriculum areas over an extended period of time. To do so, NAEP readministers materials and replicates procedures from assessment to assessment, always testing students in the same age groups (9, 13, and 17). In this manner, the long-term trends NAEP provides valuable information about progress in academic achievement and about the ability of the United States to achieve its national education goals.

To provide a numeric summary of students' performance on the assessment questions and tasks, NAEP uses a 0 to 500 scale for each subject area. Comparisons of average scale scores are provided across the years in which the NAEP long-term trend assessments have been administered and among subpopulations of students. These results chart trends from the first year in which each NAEP assessment was given: 1969/70 in science; 1971 in reading; 1973 in mathematics; and 1984 in writing.

Trends in average performance over these time periods are discussed for students at ages 9, 13, and 17 for science, mathematics, and reading. In general, the NAEP long term trends in science and mathematics show a pattern of early declines or relative stability followed by improved performance; in reading, minimal changes have occurred over the assessment period.

Science. The overall pattern of performance in science for 9-, 13-, and 17-year-olds is one of early declines followed by a period of improvement (Figure A). For 9-year-olds, the overall trend shows improvement; in 1996, the average score for these students was higher than in 1970. The overall trend for 13-year-olds was also positive, but there was no significant difference between the average science scores in 1970 and those in 1996. The average science score of 17-year-olds in 1996 was lower than the average score in 1969. Science scores have been increasing upward for all ages tested since 1982 and the publication of *A Nation at Risk*. Average scores at all three ages were higher in 1996 than in 1982 (for 17-year-olds, scores increased by 13 points; at age 13, scores increased 6 points, and at age 9, scores increased 9 points).

Mathematics. The overall pattern of mathematics achievement for 9-, 13-, and 17-year-olds shows overall improvement, with early declines or relative stability followed by increased performance (Figure B). Further, the scores of 9- and 13-year-olds were significantly higher in 1996 than in 1973. As with science, mathematics scores have also shown an upward trend at all ages since 1982 and the publication of *A Nation at Risk*. On average, the scores of 17-year-olds increased 8 points; 13-year-olds increased 5 points; and 9-year-olds increased 12 points.

Reading. The overall trend pattern in reading achievement is one of minimal changes across the assessment years (Figure C). The performance of 9-year-olds improved from 1971 to 1980, but has declined slightly since that time. However, in 1996, the average reading score for these students was higher than it was in 1971. Thirteen-year-olds showed moderate gains in reading achievement; in 1996, their average reading score was higher than that in 1971. There was an overall pattern of increase in reading scores for 17-year-olds, but the 1996 average score was not significantly different than in 1971. Reading scores have remained fairly stable between 1984 and 1996, the time period immediately following the release of *A Nation at Risk*. No significant changes at any age occurred during this time period.

Subgroup Performance on NAEP

Analyses of NAEP assessment data by race show how achievement gaps have been changing over time. In mathematics and reading, score gaps between white and black students aged 13 and 17 narrowed during the 1970s and the 1980s. Although there was some evidence of widening gaps during the late 1980s and 1990s, the score gaps in 1996 were smaller than those in the first assessment year for 13- and 17-year-olds in mathematics and for 17-year-olds in reading. Among 9-year-olds, score gaps in mathematics and reading have generally decreased across the assessment years, resulting in smaller gaps in 1996 compared to those in the first assessment year.

Since *A Nation at Risk*, performance in science has been increasing for white, black, and Hispanic students at ages 9, 13, and 17. At age 17, for example, average scores of white students increased 14 points from 1982 to 1996; for black students the increase was 25 points; and Hispanic students improved by 20 points. As a result of these increases, the gap between white and black students closed significantly (although it is still 47 points); the gap between white and Hispanic students also narrowed, though the change was not statistically significant (the gap in 1996 was 38 points).

Average mathematics scores of white, black, and Hispanic students also increased since 1982. For 17-year-olds, for example, white students improved 9 points; black students improved 14 points; and Hispanic students increased 15 points. The gaps between white and black students narrowed between 1982 and 1990, but has widened again through the 1990s, to 27 points in 1996. The gap between white and Hispanic students narrowed somewhat since 1982, though the change was not statistically significant, and the gap remained at 21 points in 1996.

Changes in reading were minimal for white, black, and Hispanic students at all ages during the years 1982 to 1996. As a result, the gaps between white and black students remained about the same (in 1996 the gap at age 17 was 29 points). The gap between white and Hispanic students also changed little (in 1996 the gap at age 17 was 30 points).

In looking at subgroup performance in NAEP, it is particularly interesting to examine how gains made by subgroups over time can be masked by simple averages. Whenever the demographic balance among subgroups shifts, it can result in what is sometimes termed “Simpson’s paradox” – which is illustrated by the NAEP long-term reading gains of 9 year-old whites, blacks, and Hispanics compared to the overall average gains shown in Figure D. Between 1971 and 1996, 9-year-old students’ average performance in reading rose by 4

points on a 500 point scale. Yet average score increases for each of the subgroups — blacks, Hispanics, and whites — exceeded the overall average increase. Why? Blacks and Hispanics, the lowest scoring subgroups represent a greater share of the total population in 1996 compared with 1971, which had the paradoxical effect of lowering overall gains even as each group's performance improved.

Framework-based Assessments in Mathematics, Reading, and Science

In addition to, and separate from, the long-term trend assessments, NAEP also provides cross sectional data based on grade level student samples. These reports, called “The Nation’s Report Card”, involve more recently developed testing instruments. Instead of repeatedly using the same sets of questions and tasks necessary to generate trend data, the Nation’s Report Card is framework-based, that is they reflect the best current thinking about what all children should know and be able to do. Each of these framework-based assessments is based on different sets of questions or tasks; therefore, the results from each cannot be directly compared.

Mathematics. The NAEP 1996 mathematics assessment continues the commitment to evaluate and report the educational progress of students at grades 4, 8, and 12. Like previous NAEP mathematics assessments in 1990 and 1992, the 1996 assessment uses a framework influenced by the Curriculum and Evaluation Standards for School Mathematics of the National Council of Teachers of Mathematics (NCTM). The 1996 framework was updated to more adequately reflect recent curricular emphases and objectives.

The framework characterizes the mathematics domain in terms of five content strands -- number sense, properties, and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and algebra and functions. Across the five content strands, the assessment examines mathematical abilities (conceptual understanding, procedural knowledge, and problem solving) and mathematical power (reasoning, connections, and communication). The positive news is that national data from the 1996 mathematics assessment showed progress in students’ mathematics performance on a broad front, as compared with both the 1990 and 1992 assessments.

- Students' scores on the NAEP mathematics scale increased for grades 4, 8, and 12 (Figure E). For all three grades scores were higher in 1996 than in 1992, and higher in 1992 than in 1990. The national average scale score for fourth-graders in 1996 was 224, an increase of 11 points from 1990; the average scale score for eighth-graders in 1996 was 272, an increase of 9 points from 1990; and the average score for twelfth-graders was 304, an increase of 10 points from 1990.
- Student performance also increased as measured by the three mathematics achievement levels set by the National Assessment Governing Board (NAGB): *Basic*, *Proficient* and *Advanced*. The percentage of students at or above the *Basic* level increased for all three grades. The percentage of fourth-grade students at or above the *Proficient* level increased between 1990 and 1992, and between 1992 and 1996, while the percentage of eighth- and twelfth-grade students at or above the *Proficient* level increased between 1990 and 1996, but was not significantly different from 1992. However, only eighth-grade students

showed an increase in the percentage at the *Advanced* level, and this increase was for the period 1990 to 1996.

- For fourth-grade students, the percentage performing at or above the *Basic* level was 64 percent in 1996, as compared to 50 percent in 1990; for eighth-grade students, 62 percent, as compared to 52 percent; and for twelfth-grade students, 69 percent, as compared to 58 percent.
- The performance of minority students, however, showed no improvement during the period, with a large performance gap persisting. For example, at grade 4 in 1996, 64 percent of black students failed to meet the *Basic* standard, in contrast to 32 percent of white students.

State data for the NAEP 1996 mathematics assessment covered fourth-graders in 47 states, territories, and other jurisdictions and eighth-graders in 44 states and jurisdictions. Many, but not all, states and jurisdictions showed increases in mathematics performance for the 1996 assessment (Figure F).

- Fourth-graders in 15 of the 39 states and jurisdictions participating in both the 1992 and 1996 assessments showed an increase in their average scale scores for 1996; 3 states showed a decrease; and 21 states had no change.
- Eighth-graders in 27 of the 32 jurisdictions participating in both the 1990 and 1996 State NAEP mathematics assessments showed an increase in their average scale scores; none declined, and 5 had no change.
- Colorado, Connecticut, Indiana, North Carolina, Tennessee, Texas, and West Virginia reported increases in the percentages of fourth-graders who scored at or above the *Basic* and *Proficient* achievement levels over the period 1992 to 1996.
- Maryland, Michigan, Minnesota, Nebraska, North Carolina, and Wisconsin reported increases in the percentages of eighth-graders that scored at or above all three achievement levels over the period 1990 to 1996.
- According to several achievement benchmarks, eighth-graders in 23 of the 32 states and jurisdictions showed improvement between 1990 and 1996. For example, the average mathematics scale scores increased in these states and jurisdictions, as well as the number of students scoring at or about the *Basic* and *Proficient* achievement levels.
- Despite these widespread increases in performance, large variations in state averages persist. The proportion of eighth-graders performing at a *Basic* or above mathematics level ranged from 36 percent in Mississippi to 77 percent in Maine and North Dakota and 78 percent in Iowa.

Reading. As is the case in mathematics, the two most recent NAEP reading assessments in 1992 and 1994 were based on a framework developed through NAGB's consensus process. The framework reflects the state of the art in curricular emphasis and objectives, as well as in assessment design. The framework defines reading in terms of three general types of text and reading situations: (1) reading for a literary experience, which focuses mostly on narrative

text; (2) reading to be informed, based on expository text; and (3) reading to perform a task, which is document based. In addition, the framework emphasizes four ways readers respond to text — they construct an initial understanding, develop an interpretation; examine the meaning to respond personally, and take a critical stance so that they might evaluate the content and/or the author’s craft.

Since 1990, the NAEP reading assessments have increasingly emphasized the importance of having students construct a response to what they have read. This has been accomplished through the use of fewer but longer text selections and an increasing number of items that require students to answer with original responses as short as one or two sentences or as long as a few paragraphs.

National data from the *NAEP 1994 Reading Report Card* showed no significant changes in average performance among the national population of fourth- or eighth-graders from 1992 to 1994. However, between these years there was a decline in the average reading performance of twelfth-graders in all three assessed purposes for reading.

- The decline in performance among twelfth graders between 1992 and 1994 was concentrated among lower performing students – those scoring at the 10th, 25th, and 50th percentiles. No significant declines were observed among twelfth graders at the 75th or 90th percentiles.
- The decline in performance among twelfth graders in 1994 also reflected in the distribution of student performance as measured against the three reading achievement levels set by the National Assessment Governing Board (NAGB). The percent of twelfth grade students who reached the *Proficient* level in reading declined from 1992 to 1994, and there was also a decrease in the percent who were at or above the *Basic* level.
- In 1994, 30 percent of fourth graders, 30 percent of eighth graders, and 36 percent of twelfth graders attained the *Proficient* level in reading. Across the three grades, 3 to 7 percent reached the *Advanced* level.
- Across the nation, there were declines in average reading performance from 1992 to 1994 for Hispanic students in grade 4, as well as for white, black, and Hispanic students in grade 12.
- Performance at all three grades was higher on average for students whose parents had more education. Among twelfth graders, the decline in average reading performance since 1992 was evident for students reporting at all levels of parental education.
- At all three grades, females had higher average reading scores than males. At twelfth grade the performance of both males and females declined between 1992 and 1994.
- In 1994, fourth, eighth, and twelfth grade students attending nonpublic schools displayed higher average reading scores than their public school counterparts. The performance of twelfth graders in public and nonpublic schools declined since 1992.

State NAEP 1992 and 1994 reading data are only available for fourth graders with 41 participating jurisdictions.

- The eight states with the highest average reading performance in 1994 among fourth graders in public schools were Maine, North Dakota, Wisconsin, New Hampshire, Massachusetts, Iowa, Connecticut, and Montana.
- Approximately 20 percent of the jurisdictions that participated in both the 1992 and 1994 reading assessments showed significant decreases in average reading performance among fourth graders. The eight jurisdictions making up the 20 percent were California, Delaware, Louisiana, New Hampshire, New Mexico, Pennsylvania, South Carolina, and Virginia.

Science. The NAEP 1996 science assessment, which gathered information about the science knowledge of the nation's fourth, eighth, and twelfth-grade students, provides baseline information about science achievement in this country. The NAEP 1996 science results are important not only because they provide this baseline information, but also because their release coincides with release of the science achievement results for the United States on the Third International Mathematics and Science Study (TIMSS). The results from these two major surveys provide valuable data on how science is taught and learned in U.S. schools.

The science framework for the 1996 NAEP science assessment was developed through a national consensus process involving educators, policymakers, science teachers, representatives of the business community, assessment and curriculum experts, and members of the general public. Two principles guide the science framework. First, the framework recognizes that scientific knowledge relies on the ability to organize disparate facts and to draw inferences from patterns and relationships. Second, the NAEP framework assumes that scientific performance depends on the ability to use scientific tools, procedures, and reasoning processes.

The core of the science framework is organized into three major fields: earth, physical, and life sciences. The assessment measures a student's ability to know and do science within these fields by testing the knowledge of important facts and concepts; the ability to explain, integrate, apply, analyze, evaluate, and communicate scientific information; and the ability to perform investigations, and evaluate and apply the results of investigations.

- Nationally, 29 percent of students in grades 4 and 8 were at or above the *Proficient* level, and 21 percent of students in grade 12 reached this level.
- Nationally, approximately 30 percent of students in grade 4 were below the *Basic* level, while nearly 40 percent of students in grades 8 and 12 failed to reach this level.
- No significant differences in percentages of male and female achievement level attainment occurred in grade 8, but at grade 4 more males than females performed at or above the *Proficient* level. At grade 12, greater percentages of males than females performed at or above the *Advanced*, *Proficient*, and *Basic* levels.
- Whites scores significantly higher than blacks and Hispanics at all three grade levels.

II. International Comparisons

International Comparisons of Mathematics and Science

NCES uses a combination of international and U.S. databases to look at the performance of our students. The combination of both types of data is required to see ourselves in stereographic or parallel perspective. U.S.-only data is blind in one eye, and international data is blind in the other. Both types of data are necessary for a clear and an accurate view of our students' performance.

TIMSS is noteworthy not only because of its scope and magnitude, but also because of innovations in its design. In this international study, NCES along with the National Science Foundation (NSF) combined multiple methodologies to create an information base that goes beyond simple student test score comparisons to examine the fundamental elements of schooling. Innovative research techniques include analyses of textbooks and curricula, videotapes, and ethnographic case studies. The result is a more complete portrait of how U.S. mathematics and science education differs from that of other nations, especially in extended comparisons with Germany and Japan.

The information in these reports can serve as a starting point for our efforts to define a "world-class" education. If the United States is to improve the mathematics and science education of its students, we must carefully examine not just how other countries rank, but also how their policies and practices help students achieve. TIMSS shows us where U.S. education stands -- not just in terms of test scores, but also what is included in textbooks, taught in the schools, and learned by students. Examining these data provides a valuable opportunity to shed new light on education in the United States through the prism of other countries. At the same time, we should avoid the temptation to zero in on any one finding or leap to a conclusion without carefully considering the broader context.

Our students' international standing declines as students progress through school, according to TIMSS. Overall, U.S. fourth-graders scored above the international average in both science and mathematics. Our eighth-graders scored above the international average in science but below it in mathematics. In twelfth-grade, the scores of both our overall student population tested on general mathematics and science knowledge, and of our more advanced students tested in mathematics and physics, were well below the international average.

Fourth-Grade Findings. In both mathematics and science, U.S. fourth-graders performed above the international average. In mathematics, of the 26 participating TIMSS countries, U.S. fourth-graders outperformed students in 12 countries and were outperformed by students in seven countries. In science, U.S. students outperformed students in 19 countries, and were outperformed by students in only one country—Korea. In the six mathematics content areas, U.S. fourth-graders exceeded the international average in five. In the science content areas, U.S. fourth-graders exceeded the international average in all four areas assessed.

Eighth-Grade Findings. Data on eighth-grade performance from TIMSS suggests a general improvement in U.S. eighth-grade science scores as compared to a prior 1991 international assessment that placed U.S. students below average, though the tests and the set of participating nations have changed. The TIMSS data, however, show that U.S. eighth-grade

students' mathematics performance remains slightly below the international average. U.S. eighth-grade students scored lower, on average, in mathematics than students in Canada, France, and Japan, and scored about the same as students in England and Germany. In science, eighth grade students from the United States scored higher, on average, than students in France, about the same as students in Canada, England, and Germany, and lower than students in Japan. Figure G summarizes U.S. performance by content area on the fourth- and eighth-grade assessments.

Twelfth-Grade Findings. The twelfth-grade TIMSS included 21 countries that conducted assessments of their students' general knowledge in mathematics and science during their last year in secondary school. Japan and other Asian countries that traditionally perform well in mathematics and science did not participate in the twelfth-grade TIMSS. Even with those Asian countries excluded, the United States performed relatively poorly. In the mathematics general knowledge assessment, U.S. twelfth-grade students were outperformed by 14 countries, and outperformed two countries. U.S. students performed the same as students in four other countries. In science, U.S. twelfth-grade students were outperformed by students in 11 countries, and outperformed students in two countries. U.S. students performed the same as students in seven other countries (Figure H).

Average test scores can mask important differences in the distribution of scores. For example, as a result of our country's diverse population, U.S. test score averages could be unduly lowered by a relatively large group of low-scoring students. In the twelfth-grade TIMSS assessments, however, the distribution of scores among U.S. students was no wider than that in most other participating countries; the U.S. scores also start and end lower than those in higher scoring countries. We also like to think that at least America's "best and brightest" students are among the smartest in the world; again, TIMSS findings suggest otherwise. Sixteen countries assessed advanced mathematics and physics among a select group of advanced students. In advanced mathematics, 11 countries outperformed the U.S., and no countries performed more poorly. In physics, 14 countries outperformed the U.S.; again, no countries performed more poorly (Figure I).

Several other factors suggested by observers also do not account for the relatively poor performance of U.S. students in grade 12. For example, it is not the case that a greater proportion of U.S. students complete secondary school than in most of the other countries participating in this phase of TIMSS. Thus, the vast majority of U.S. young people are not being compared only to an elite in other countries. Furthermore, in TIMSS, the general pattern was that countries with higher proportions of young people enrolled in and completing secondary school outperformed countries with lower proportions. The decentralized nature of decision-making about curriculum did not explain the poor performance of U.S. students. Some countries with decentralized decision-making outperformed us and some did not. The same was true of countries with centralized decision-making. Finally, while U.S. students on average were about a half a year younger than the average for all 21 countries, the age differential is not a major factor contributing to our poor performance. Not only is the age differential relatively small (and it is even less in the advanced assessments), countries in which the average age of the students was similar to or younger than the U.S. also outperformed us.

Among the other achievement findings drawn from the TIMSS:

- In comparison with their international counterparts, U.S. students performed better in science than in mathematics at all three grade levels;
- Among U.S. students, there is no significant gender gap in mathematics at any grade level for the general populations tested. However, in fourth-grade and twelfth-grade science, and in twelfth-grade advanced mathematics and physics, male students performed better than female students.
- At grade 4, 16 percent of U.S. students would be in the international top 10 percent in science; at grade 8, 13 percent;
- In mathematics, 9 percent of U.S. fourth-graders would be in the international top 10 percent in mathematics, compared to 5 percent of eighth-graders.

Lessons From TIMSS

While TIMSS has given us information on our international standing, it is most valuable in telling us what factors are related to high achievement in schools. The overarching message is that there is no easy solution or single nostrum that will magically increase our nation's performance. Indeed, TIMSS shows us that many of the cure-alls recommended in the past are not associated with high performance in all nations. For example, more seat time in math and science, more homework, and less television have often been recommended as methods for increasing student performance.

These strategies may indeed be effective in the case of individual students or schools, yet TIMSS has shown us another perspective. Comparisons of eighth-grade students, teachers, and classrooms in the U.S., Japan, and Germany have been particularly revealing. For example, U.S. eighth graders already spend more seat time in math and science classes than students in Japan and Germany. Japan outperforms us at this grade level, while Germany does not, so this shows that more seat time is not necessarily a magic tonic. With respect to homework, U.S. eighth-grade teachers already assign more homework, spend more class time discussing it, and are more likely to count it toward grades than teachers in Japan. Japanese eighth graders also watch just as much TV as students in the U.S. The most recent TIMSS also found that the relatively poor performance of U.S. twelfth-grade students is not related to hours spent on homework, the use of calculators or computers, time spent watching television or working at a paid job, or to attitudes toward mathematics and science.

These and other TIMSS findings show us that there is no single easy answer to achieving high performance in mathematics and science. But the TIMSS and other NCEs data sources do suggest some problems in U.S. mathematics and science education that may help explain our relatively low achievement at the higher grade levels. These data suggest that three issues are worth our attention: curriculum, coursetaking, and teacher preparation.

First, both the mathematics and science curricula in American high schools have been criticized for their lack of coherence, depth, and continuity—for covering too many topics at

the expense of in-depth understanding. As a result, our secondary school curricula leave American students with a more limited opportunity to learn than their counterparts have in other countries. For example, while most other countries introduce algebra and geometry in the middle grades, in the U.S. only 25 percent of students take algebra before high school. The TIMSS also demonstrated the relative “slowness” of our curricula. The study found that the topics on the twelfth- grade general knowledge mathematics assessment were covered by the ninth grade in the U.S., but by the seventh grade in most other countries. The topics on the general science assessment were covered by the eleventh grade in the U.S., but by the ninth grade in most other countries.

Students’ exposure to challenging mathematics and science content is further limited by their coursetaking behavior. Despite some recent increases in academic coursetaking, fully 90 percent of all U.S. high school students stop taking mathematics before getting to calculus. Even among college-bound seniors, 52 percent have not taken physics, 48 percent have not taken trigonometry, and 77 percent have not taken calculus; almost one-third (31 percent) had not taken four years of mathematics. Among 1994 high school graduates, only 9 percent had taken calculus and 24 percent had taken physics.

Finally, courses and curricula do not teach themselves. At the most basic level, the education system relies on knowledgeable, well-trained teachers to convey the information students need to learn. What teachers do not know, they cannot teach. And our data suggest that considerable percentages of our mathematics and science teachers have not been adequately exposed to the information they teach. Figure J shows that in 1993-94, 28 percent of public high school (grade 9-12) mathematics teachers and 18 percent of public high school science teachers were teaching out-of-field (that is, without a major or minor in their subject). Within science sub-fields, 31 percent of life science (biological/life sciences) teachers and 55 percent of physical science (chemistry, physics, earth science, and physical science) teachers lacked a major or minor in their sub-field. In addition, 24 percent of mathematics teachers and 17 percent of science teachers lacked state certification in their teaching field.

In short, TIMSS does dispel myths, but more importantly, it shows us our own education system in clearer perspective. In our quest for factors related to better student performance, TIMSS encourages us to focus on rigorous content, focused curriculum, good teaching, and good training for teachers. TIMSS has shown us that the typical U.S. eighth-grade mathematics class usually discusses material taught at the seventh-grade around the world. Compared to those in Japan, our mathematics teachers tend to focus on teaching specific math skills, rather than higher-level mathematical problem solving. For example, U.S. eighth-grade math teachers are more likely to merely state rather than explain mathematical concepts. Further, our curriculum includes more topics, and our teachers are more frequently interrupted by loudspeakers and other outside agents, while they are teaching than are teachers in Japan and Germany. Our teachers also lack a one or two year apprenticeship in teaching before they become teachers, as is the case in these two other countries. Clearly TIMSS shows us that while it may not be easy, important change is needed to help our nation continue to improve its performance.

International Comparisons of Reading

In 1991, the IEA Reading Literacy Study assessed the reading literacy of fourth-graders (in 27 countries) and ninth-graders (in 31 countries). The underlying framework for this assessment paralleled the NAEP framework in that it too defined reading in terms of three text types – narrative, expository and document. In contrast to the NAEP Reading Report Card, this study painted a more positive picture of the reading literacy of American students.

- American fourth-graders were outperformed only by Finland; U.S. students performed about the same as students from Sweden, while outperforming students from 24 other nations.
- American ninth-graders' performance was equivalent to that of students from 15 other nations; Americans outperformed students from 14 nations, while only the students from Finland did better than our students.

Considering only those countries that were then part of the Organization for Economic Cooperation and Development (OECD), the study's findings indicate that:

- Among fourth-graders the reading performance of about 60 percent of U.S. students meets or exceeds the OECD average on two scales – narrative (which corresponds roughly with the NAEP “reading for a literary experience” scale), and expository (which corresponds roughly with the NAEP “reading to be informed” scale). About 70 percent of American fourth graders meet or exceed the OECD average on the third IEA reading scale – documents (which roughly corresponds with the NAEP “reading to perform a task” scale).
- The comparative advantage of American ninth-grade students was not as great as that of the fourth graders. Between 52 and 55 percent of U.S. ninth-graders met or exceeded the OECD average on the three scales.
- Most groups of American students, even the most disadvantaged, outperform the OECD reading average, with only a few exceptions – black students in both grades and students in 9th grade whose parents did not complete high school do not consistently meet or exceed the OECD average.

The difference between the NAEP view of America's fourth, eighth, and twelfth-grade students' reading proficiency and that emerging from the IEA data may be attributed to two very important differences in these assessments. First, there are distinct differences in the way that the data are benchmarked. IEA reporting is based on comparisons of student performance across countries, while much of NAEP reporting is based on student performance against a desired standard defined by NAGB. Second, the IEA test mainly asks students to recognize details and to make simple inferences and literal interpretations while the NAEP test goes further, i.e., requiring students to identify themes to detect the author's point of view, to make larger inferences, and to state a position with supporting citations from the text. These differences in benchmarking and in test rigor raise important questions. Primarily, we must consider what benchmarks are reasonable for our society. One way to

examine this issue is to look at achievement or proficiency data in relation to important outcome measures, as we will discuss next.

International Perspective on Labor Force Proficiency

Literacy has been viewed as one of the fundamental tools necessary for successful economic performance in industrialized societies. As society becomes more complex and low-skill jobs continue to disappear, concern about adults' ability to use written information to function in society continues to increase. Within countries, literacy levels are affected both by the quality and quantity of the population's formal education, as well as by participation in informal learning activities.

The most recent international adult literacy data (1996) demonstrate that the U.S. appears most similar to New Zealand and the United Kingdom in the overall distribution of literacy skills. (See Figure K.) These three countries had close to 20 percent of their adult population at both the high and low ends of the literacy scale (Level 1 and Levels 4 and 5). In contrast, the performance of our European counterparts was concentrated in the middle literacy levels, with at least two-thirds of the adult population in the Netherlands, Switzerland (both French and German speaking) and Germany at Literacy Levels 2 or 3. While Sweden tended to have the greatest concentration at the higher end of the scale, Poland's adults were concentrated at the lower end.

In the United States, as you might expect, workers with higher adult literacy scores are unemployed less and earn more than workers with lower literacy scores. Unemployment rates are especially high for workers in the two lowest levels of literacy—levels 1 and 2—on each of the three literacy scales. For these workers, the unemployment rate ranges from 12 percent for workers with level 2 quantitative literacy to nearly 20 percent for those with level 1. Unemployment rates for individuals in the two highest literacy levels—levels 4 and 5—are less than 6 percent.

Workers with high literacy scores earn more than other workers do, on average (Figure L). On the prose scale, for example, full-time workers in level 3 earn a mean weekly wage 50 percent higher than that of their counterparts in level 1. Those in level 5 earn a weekly wage 71 percent higher than the wage of those in level 3. Thus, academic skills do make a difference in both earnings and employability.

III. Changes in Student Behavior Since *A Nation at Risk*

In addition to reviewing changes in student achievement since *A Nation at Risk*, as well as our comparative international educational standing, it is instructive to look at other significant changes in the educational landscape since the publication of this seminal work. Three are especially worthy of note: the decline in the high school dropout rate, an increase in the educational aspirations and college attendance rates of high school seniors, and increases in the academic course load of high school students. Each of these changes indicate noteworthy positive responses to what was called for in *A Nation at Risk*.

Dropping Out of School

There has been a reduction in the drop out rate since *A Nation at Risk*. Most of this decline occurred during the 1980s, and was especially pronounced for blacks. Over the last decade, 300,000 to 500,000 students in grades 10 through 12 left school each year without successfully completing a high school program. In October 1996, nearly 3.6 million 16- to 24-year-old youth were not enrolled in a high school program and lacked a high school credential. These young adults accounted for 11 percent of the 32 million 16- to 24-year-olds in the United States. Nevertheless, this 1996 dropout rate was three points lower than the 1982 dropout rate of 14 percent. And, the dropout rate for Black youth during this period fell from 18 to 13 percent.

The dropout rates of 16-to-24-year-olds Hispanics remained at levels substantially higher than the dropout rates experienced by their white and black peers (Figure M). And, in contrast to the decline among black and white 16-to-24-year olds, the dropout rates for Hispanics has not changed significantly since 1982. In 1996, 29 percent of Hispanics were not enrolled in school and had not completed high school; however this percentage includes young immigrants who came to the United States without high school credentials and never enrolled in a U.S. school. The dropout rate for Hispanic immigrants aged 16- to 24-years-old was 44 percent, compared to the dropout rate for first-generation Hispanics born in the United States, which was 17 percent (Figure N).

Educational Aspirations and College Attendance

One of the most dramatic changes taking place since *A Nation at Risk* is that the hopes of high school seniors for the future increasingly include more education. In 1992, 69 percent of seniors said that they hoped to graduate from college, compared with 39 percent of 1982 seniors. Moreover, 33 percent said they hoped to earn a postgraduate degree as compared with 18 percent in 1982. The proportion of minority students aspiring to postgraduate degrees was about the same, or higher, than for whites. Not surprisingly, these higher student aspirations have been accompanied by substantial increases in actual college attendance. The proportion of high school graduates going directly on to college rose from 51 percent in 1982 to 65 percent in 1996.

Coursetaking Patterns in High School

One of the important elements in the recommendations in *A Nation at Risk* was to increase the academic course load of high school students. Since the release of that report, most states have raised course requirements for high school graduation and most states have mandated student-testing standards. As a result, both college-bound and non-college-bound students now take more academic courses than their counterparts did a decade before. In 1982, the average high school graduate completed 2.6 Carnegie units in mathematics and 2.2 units in science. By 1994, the average number of Carnegie units completed had risen to 3.4 in mathematics, and 3.0 units in science. Foreign language units rose from 1.0 to 1.8, and coursework in English and social studies also increased. The increase in the average units completed means that more students are now taking advanced mathematics courses, such as calculus, which was completed by 9 percent of the 1994 graduates compared to 5 percent of the 1982 graduates. Similarly, the proportion of graduates completing a physics course rose from 14 percent in 1982 to 24 percent in 1994 (Figure O).

A Nation at Risk recommended that high school students complete a “New Basics” curriculum that included a minimum number of courses in the core academic areas of English (4), Mathematics (3), Science (3), and Social Studies (3). Since the release of these “New Basics” recommendations, high school graduates have taken more courses overall, particularly academic courses. The proportion of students completing the “New Basics” core curriculum in English, mathematics, science, and social studies has increased; and greater percentages are taking Advanced Placement (AP) courses. In 1982, 14 percent of high school graduates earned the credits recommended in *A Nation at Risk*; by 1994, 50 percent had done so. The percentage of graduates who have completed the more extensive recommendations for college-bound students, which include the “New Basics,” plus 2 years of foreign language instruction and a half-year of computer science, rose from 2 percent in 1982 to 25 percent in 1994.

Even though we cannot establish a cause and effect relationship, it is interesting to compare the average mathematics and science performance of 17-year-olds, as measured by our National Assessment of Education Progress, and the increase in course taking. The mathematics performance of 17-year-olds rose by 7 points between 1982 and 1994, which roughly equates to about 2/3 of the typical grade to grade progress. This increase compares closely to the rise of .8 average units of mathematics completed by high school graduates. The science performance of 17-year-olds rose by 11 points between 1982 and 1994, compared to an average increase of .9 science units completed by high school graduates.

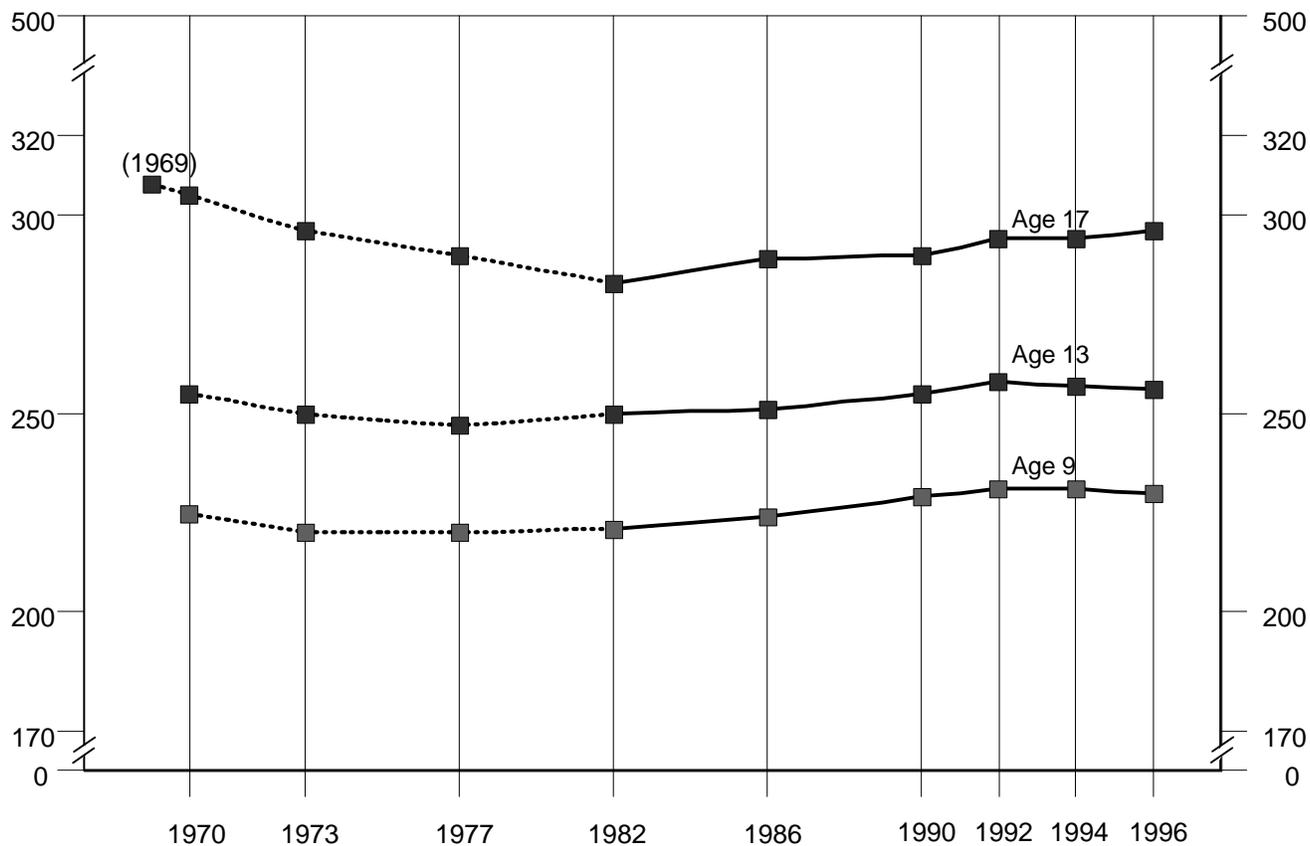
Conclusion

Whatever else one might argue is the legacy of *A Nation at Risk*, it clearly signaled the recognition of *educational performance* as a national concern, an issue of national importance. In times like this, Federal statistical agencies, such as the National Center for Education Statistics, play a critical role.

First, they provide the data that researchers and statistical analysts need. As demonstrated by the numbers presented here, there are large differences in how well students do -- across time, across countries, and sometimes across groups. It falls typically to researchers to untangle these relationships, to separate educational inputs from outputs, and to identify the processes that contribute most powerfully to student performance. Second, statistical agencies aid policymakers in a more direct manner -- by sounding alarms when problems arise or issues emerge that deserve public attention. Informing policymakers with data that clarify where problems exist and what issues are most pressing is one of the Federal government’s most vital roles.

Figure A. Trends in Average Scale Scores for the Nation: 1969–70 to 1996

SCIENCE

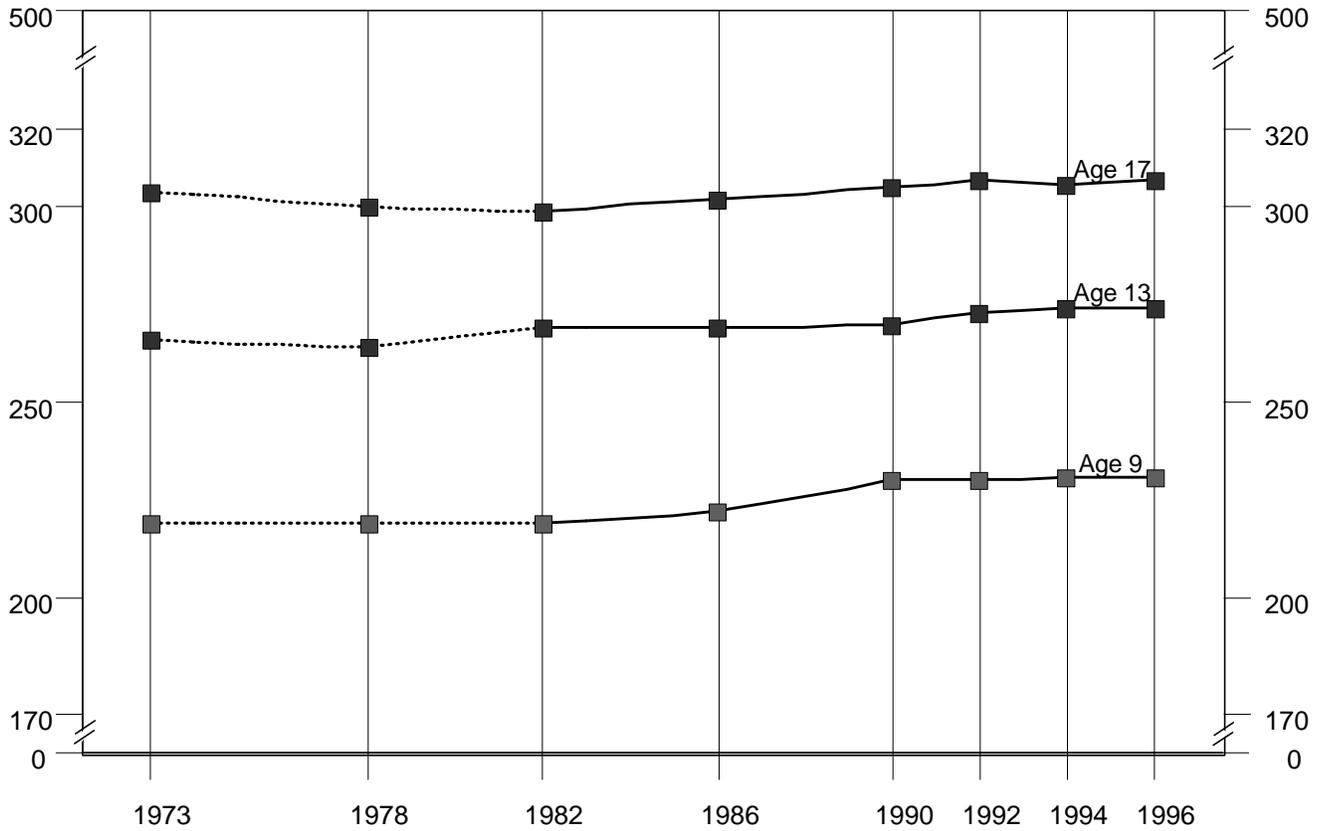


	1970	1973	1977	1982	1986	1990	1992	1994	1996
■ Age 17	305	296	290	283	289	290	294	294	296
■ Age 13	255	250	247	250	251	255	258	257	256
■ Age 9	225	220	220	221	224	229	231	231	230

SOURCE: National Assessment of Educational Progress (NAEP), *1996 Long-Term Trend Assessment*.

Figure B. Trends in Average Scale Scores for the Nation: 1973 to 1996

MATHEMATICS



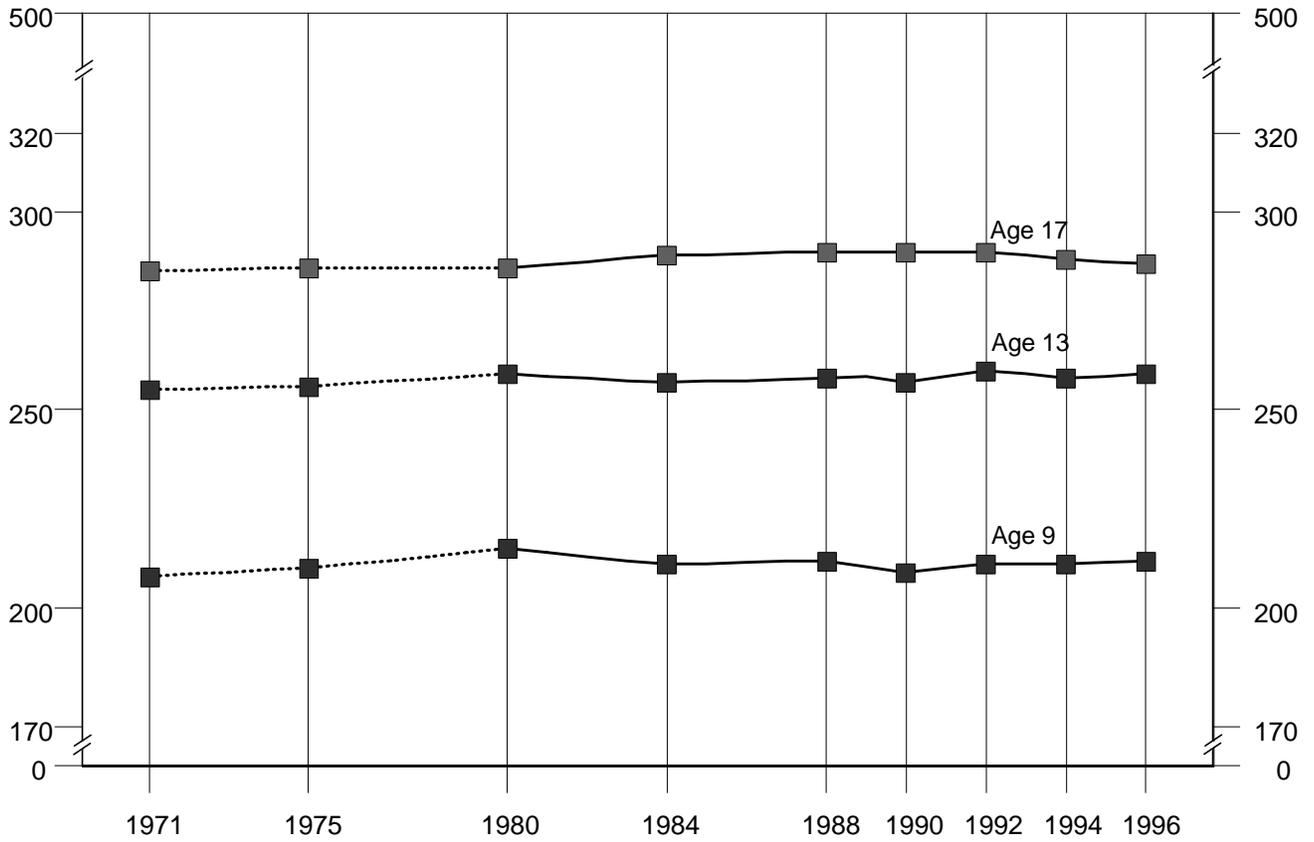
	1973	1978	1982	1986	1990	1992	1994	1996
--	------	------	------	------	------	------	------	------

■ Age 17	304	300	299	302	305	307	306	307
■ Age 13	266	264	269	269	270	273	274	274
■ Age 9	219	219	219	222	230	230	231	231

SOURCE: National Assessment of Educational Progress (NAEP), 1996 Long-Term Trend Assessment.

Figure C. Trends in Average Scale Scores for the Nation: 1971 to 1996

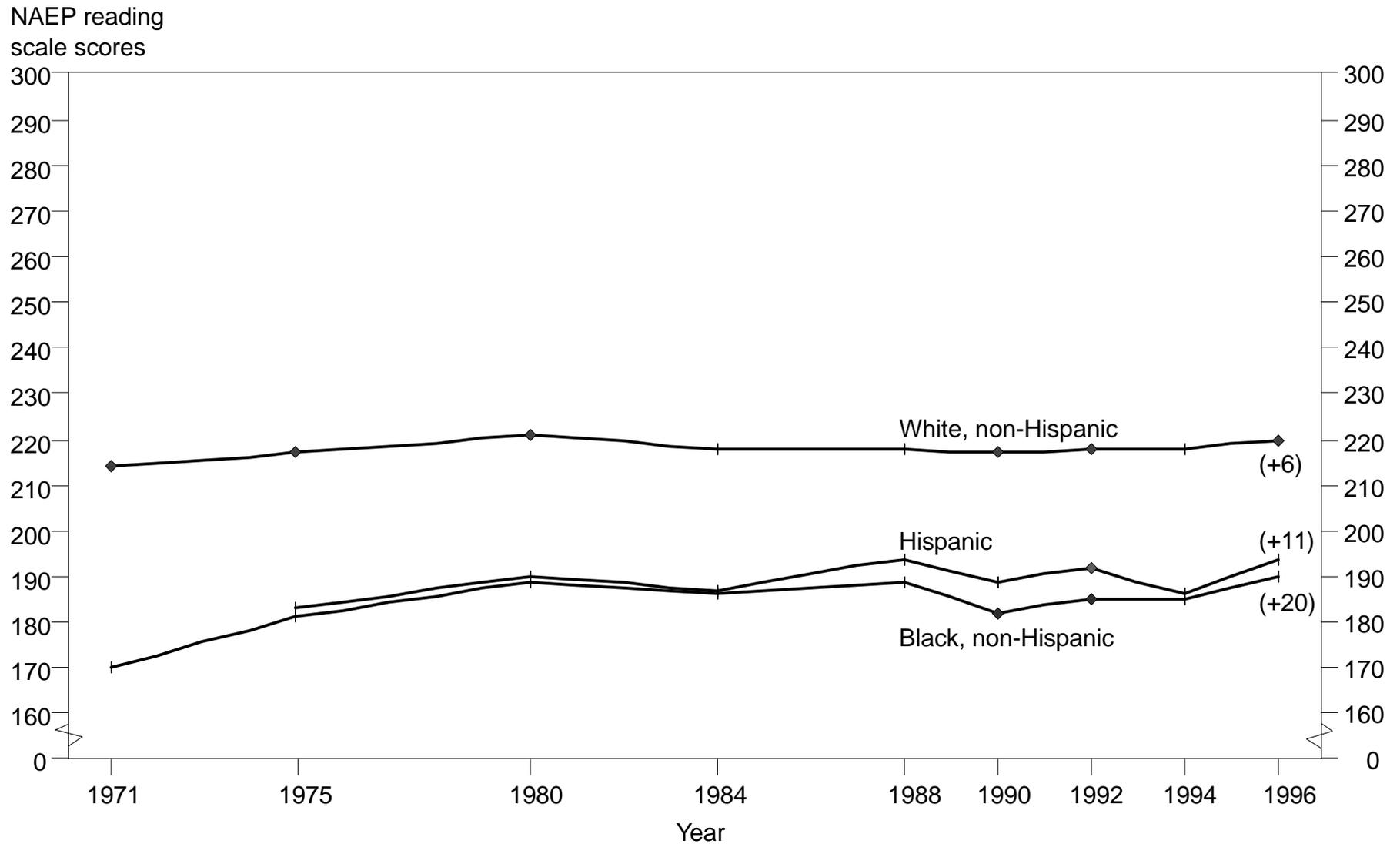
READING



	1971	1975	1980	1984	1988	1990	1992	1994	1996
■ Age 17	285	286	286	289	290	290	290	288	287
■ Age 13	255	256	259	257	258	257	260	258	259
■ Age 9	208	210	215	211	212	209	211	211	212

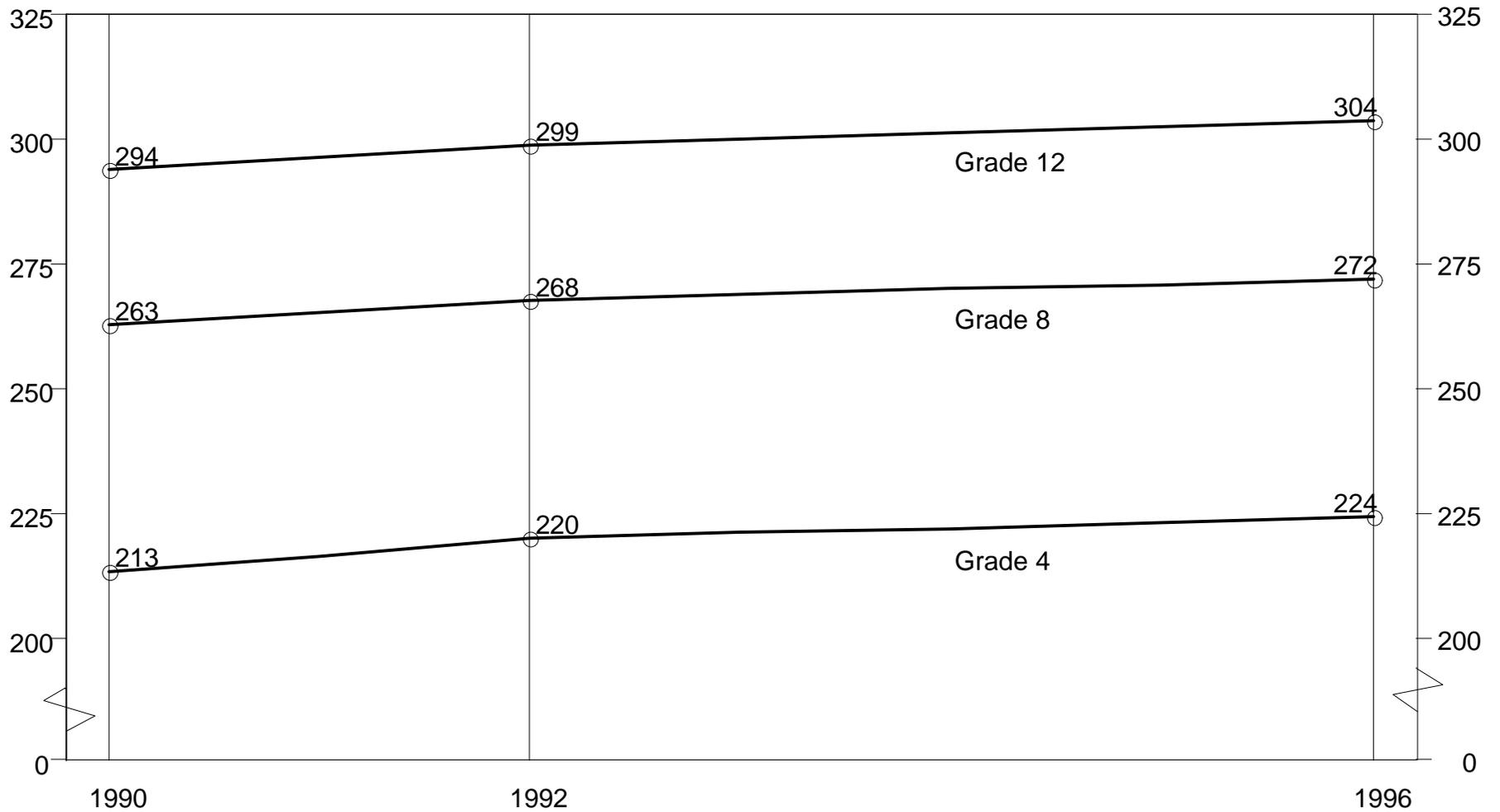
SOURCE: National Assessment of Educational Progress (NAEP), 1996 Long-Term Trend Assessment.

Figure D. Trends in Average Reading Scale Scores for 9-year-olds, by Race/ethnicity: 1971–1996



SOURCE: U.S. Department of Education, National Center for Education Statistics, Office of Educational Research and Improvement, *NAEP 1996 Trends in Academic Progress*.

Figure E. Average Mathematics Scale Scores: 1990, 1992, and 1996



* Indicates a significant difference from 1990.

+ Indicates a significant difference from 1992.

Figure F. State Improvement in Mathematics Achievement

States Showing Statistically Significant Improvement in Average Mathematics Performance, and in the Percentage of Students Performing At or Above Two Designated Achievement Levels (Both "Basic" and "Proficient")

Grade 4 (1992 to 1996): 7 states

Colorado
Connecticut
Indiana
North Carolina
Tennessee
Texas
West Virginia

Grade 8 (1992 to 1996): 3 states

Michigan
North Carolina
West Virginia

Grade 8 (1990 to 1996): 23 states

Arkansas	Minnesota
Arizona	Nebraska
California	New Mexico
Colorado	New York
Delaware	North Carolina
Florida	Oregon
Kentucky	Rhode Island
Hawaii	Texas
Indiana	West Virginia
Iowa	Wisconsin
Maryland	Wyoming
Michigan	

Figure G. U.S. Mathematics and Science Performance at a Glance

How did U.S. Students Compare with the international Average in...?	At Grade 4? (26 nations)	At Grade 8? (41 nations)
Mathematics overall	Above	Below
Science overall	Above	Above
<u>Mathematics Content Areas:</u>		
Data representation, analysis, and probability	Above	Above
Geometry	Above	Below
Whole numbers	Above	X
Fractions and proportionally	Above	X
Patterns, relations, and functions	Above	X
Measurement, estimation, and number sense	Below	X
Fractions and number sense	X	Same
Algebra	X	Same
Measurement	X	Below
Proportionality	X	Below
<u>Science Content Areas:</u>		
Earth Science	Above	Above
Life Science	Above	Above
Environmental issues and the nature of science	Above	Above
Physical science	Above	X
Chemistry	X	Same
Physics	X	Same
<u>What Percentage of U.S. Students Would Be in the International Top Ten Percent In...?</u>		
	At Grade 4?	At Grade 8?
Mathematics	9%	5%
Science	16%	13%

Above = U.S. average performance is higher than the average of participating nations at that grade.

Below = U.S. average performance is lower than the average of participating nations at that grade.

Same = U.S. average performance not significantly different than the average of participating nations at that grade.

X = Separate content area score not reported for this grade level.

SOURCE: Third International Mathematics and Science Study, *Pursuing Excellence: A Study of U.S. Twelfth-Grade Mathematics and Science Achievement in International Context*.

Figure H.

Average Mathematics General Knowledge Performance

Nations with average scores significantly higher than the U.S.	
Nation	Average
(Netherlands)	560
Sweden	552
(Denmark)	547
Switzerland	540
(Iceland)	534
(Norway)	528
(France)	523
New Zealand	522
(Australia)	522
(Canada)	519
(Austria)	518
(Slovenia)	512
(Germany)	495
Hungary	483

Nations with average scores not significantly different from the U.S.	
Nation	Average
(Italy)	476
(Russian Federation)	471
(Lithuania)	469
Czech Republic	466
(United States)	461

Nations with average scores significantly lower than the U.S.	
Nation	Average
(Cyprus)	446
(South Africa)	356

International Average = 500

Average Science General Knowledge Performance

Nations with average scores significantly higher than the U.S.	
Nation	Average
Sweden	559
(Netherlands)	558
(Iceland)	549
(Norway)	544
(Canada)	532
New Zealand	529
(Australia)	527
Switzerland	523
(Austria)	520
(Slovenia)	517
(Denmark)	509

Nations with average scores not significantly different from the U.S.	
Nation	Average
(Germany)	497
(France)	487
Czech Republic	487
(Russian Federation)	481
(United States)	480
(Italy)	475
Hungary	471
(Lithuania)	461

Nations with average scores significantly lower than the U.S.	
Nation	Average
(Cyprus)	448
(South Africa)	349

International Average = 500

NOTE: Nations not meeting international sampling and other guidelines are shown in parentheses.
 SOURCE: Third International Mathematics and Science Study, *Pursuing Excellence: A Study of U.S. Twelfth-Grade Mathematics and Science Achievement in International Context.*

Figure I.

**Average Advanced Mathematics Performance of
Advanced Mathematics Students**

Nations with average scores significantly higher than the U.S.	
Nation	Average
France	557
(Russian Federation)	542
Switzerland	533
(Australia)	525
(Denmark)	522
(Cyprus)	518
(Lithuania)	516
Greece	513
Sweden	512
Canada	509
(Slovenia)	475

Nations with average scores not significantly different from the U.S.	
Nation	Average
(Italy)	474
Czech Republic	469
(Germany)	465
(United States)	442
(Austria)	436

Nations with average scores significantly lower than the U.S.	
Nation	Average
None	

International Average = 501

Average Physics Performance of Advanced Science Students

Nations with average scores significantly higher than the U.S.	
Nation	Average
Norway	581
Sweden	573
(Russian Federation)	545
(Denmark)	534
(Slovenia)	523
(Germany)	522
(Australia)	518
(Cyprus)	494
(Latvia)	488
Switzerland	488
Greece	486
(Canada)	485
France	466
Czech Republic	451

Nations with average scores not significantly different from the U.S.	
Nation	Average
(Austria)	435
(United States)	423

Nations with average scores significantly lower than the U.S.	
Nation	Average
None	

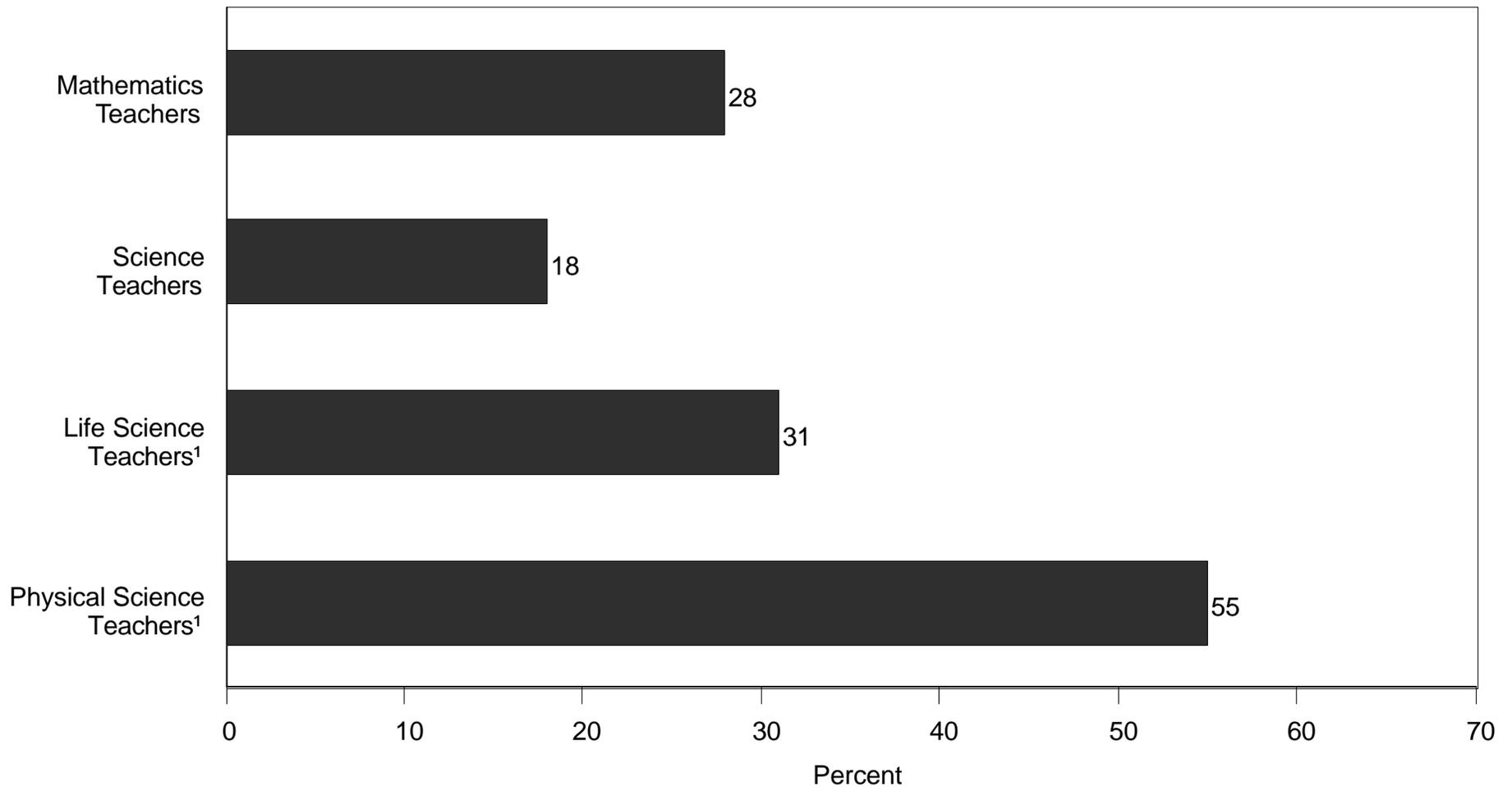
International Average = 501

NOTE: Nations not meeting international sampling and other guidelines are shown in parentheses.

SOURCE: Third International Mathematics and Science Study, *Pursuing Excellence: A Study of U.S.*

Twelfth-Grade Mathematics and Science Achievement in International Context.

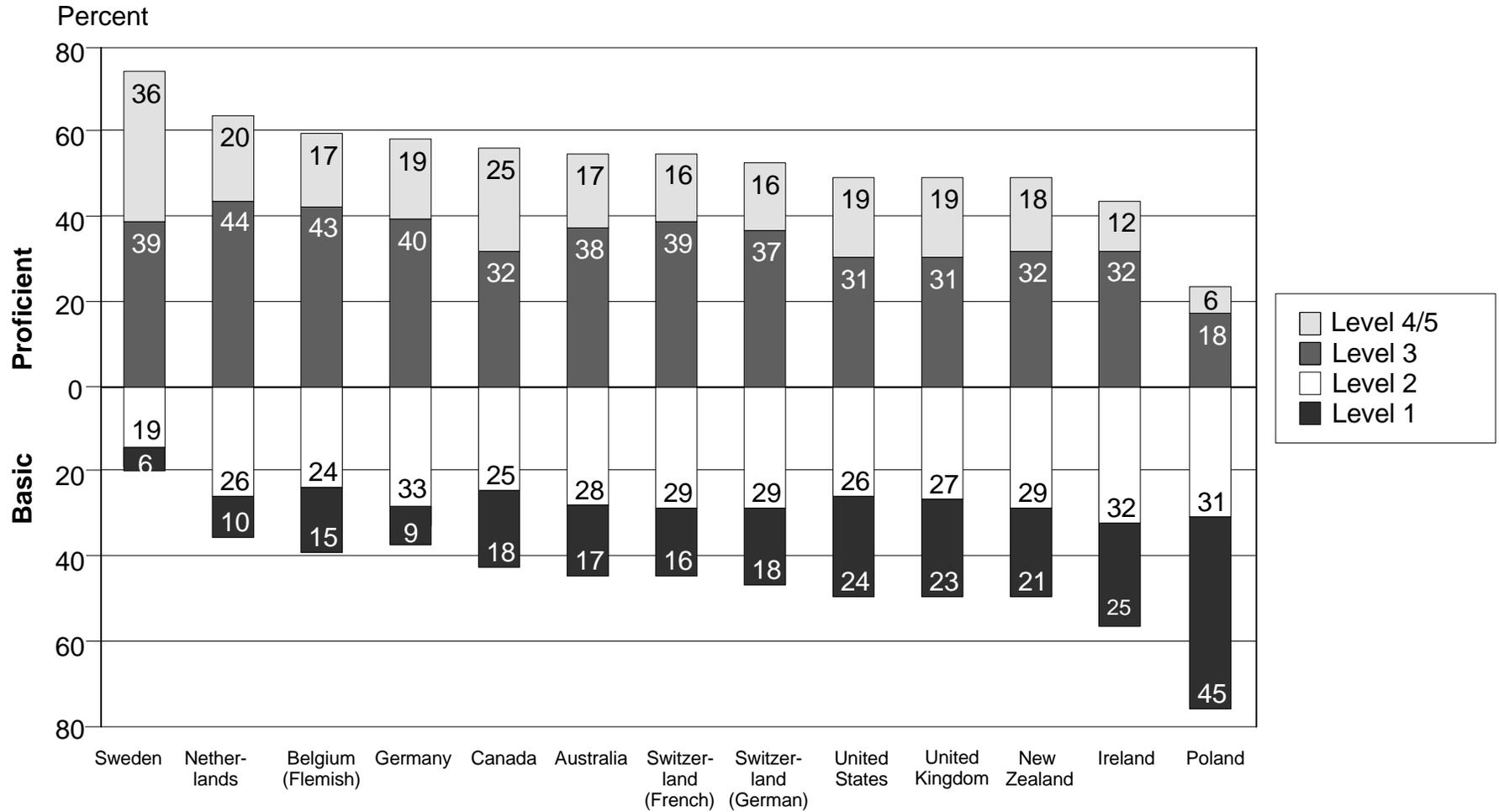
Figure J. Percentage of Public High School (grade 9–12) Mathematics and Science Teachers Without a Major or Minor in Their Field



¹ These percentages represent teachers without a major or minor in their respective science sub-field.

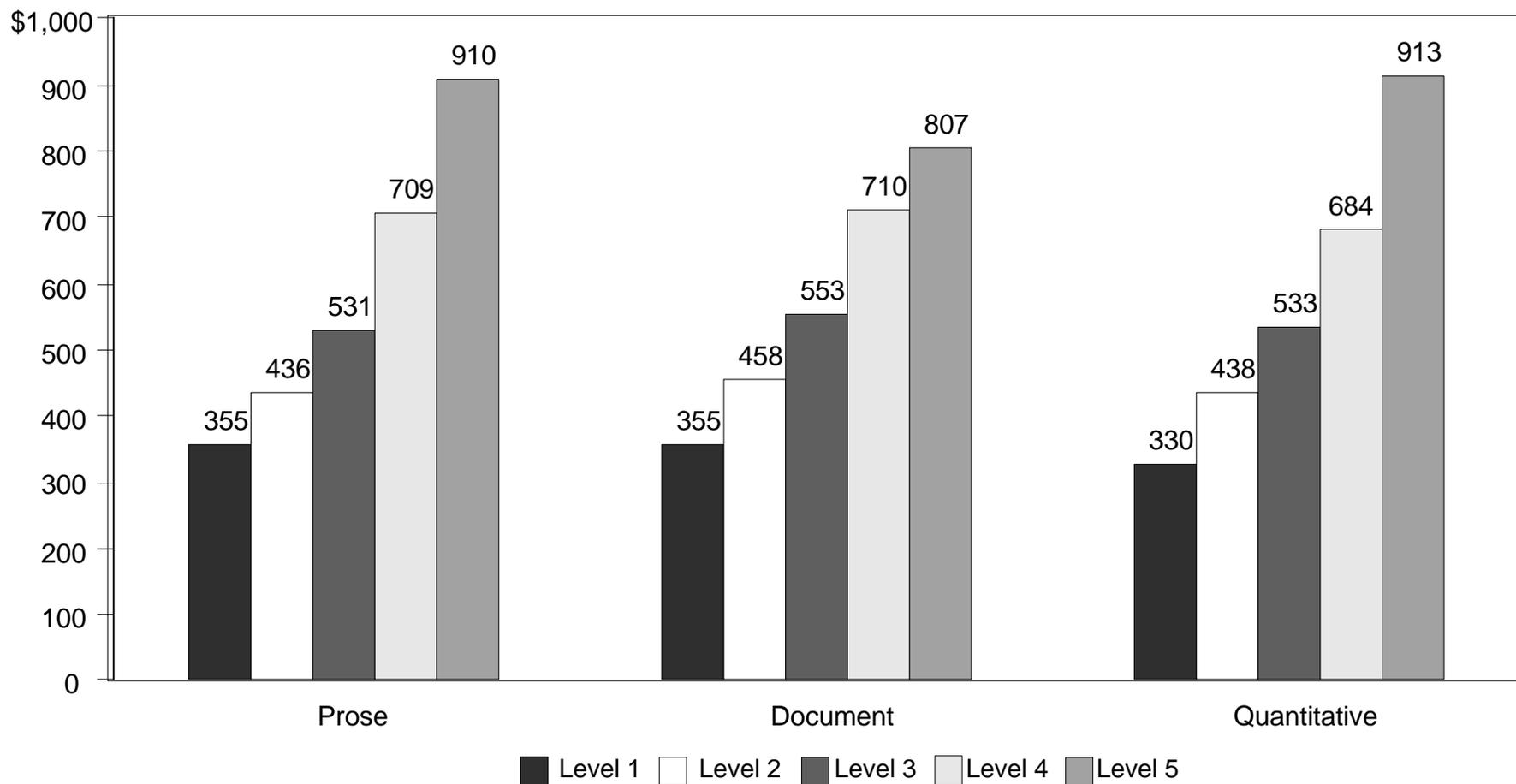
SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey, 1993–94 (Teacher Questionnaire).

Figure K. National Document Literacy Levels, Percentage of Adult Population Age 16 to 65, All Nations: 1994 and 1996



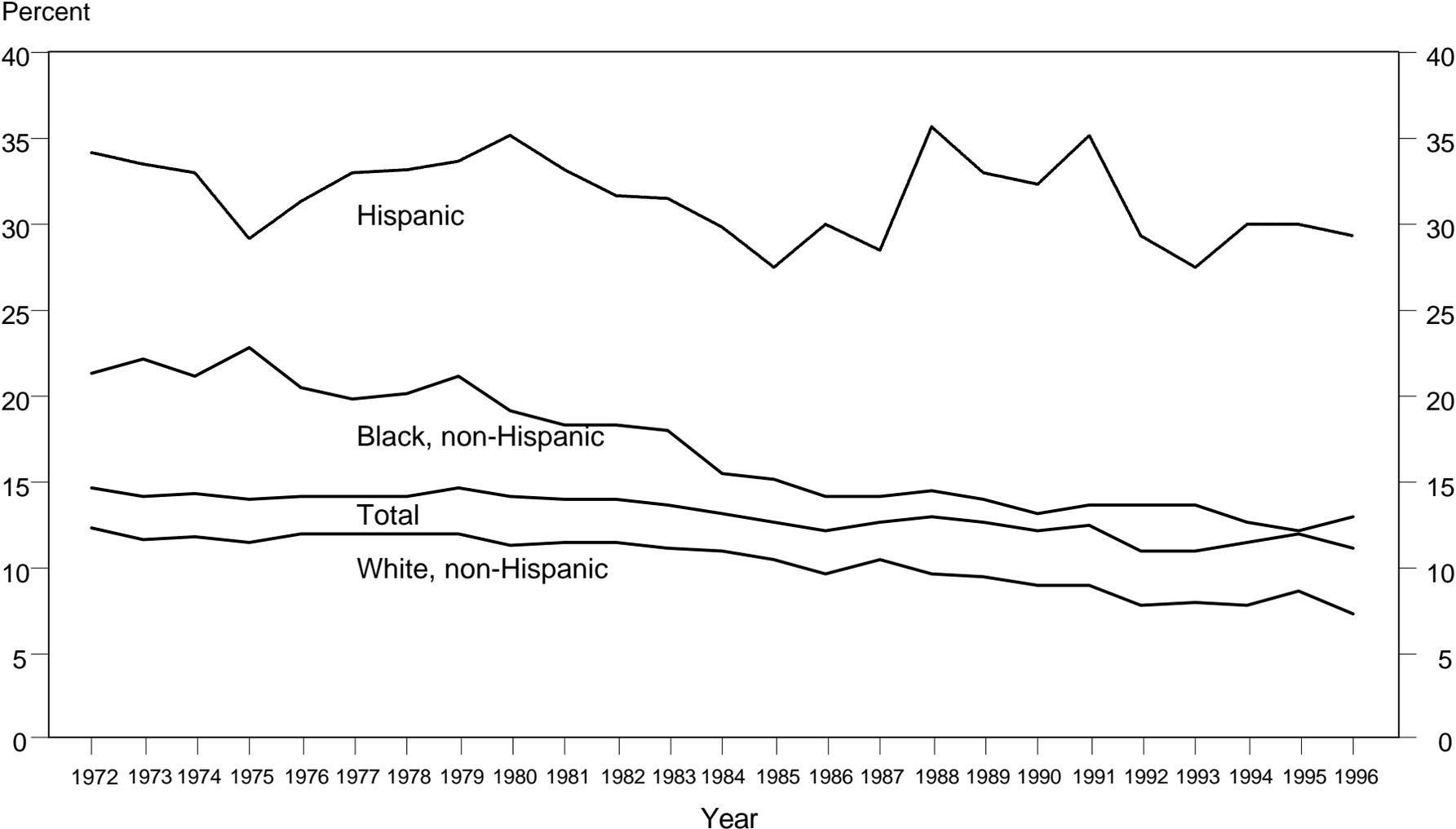
SOURCE: Organization for Economic Cooperation and Development and Statistics Canada.

Figure L. Mean Weekly Earnings of Full-time Workers, by Proficiency Level on Three Literacy Scales: 1992



SOURCE: National Adult Literacy Survey.

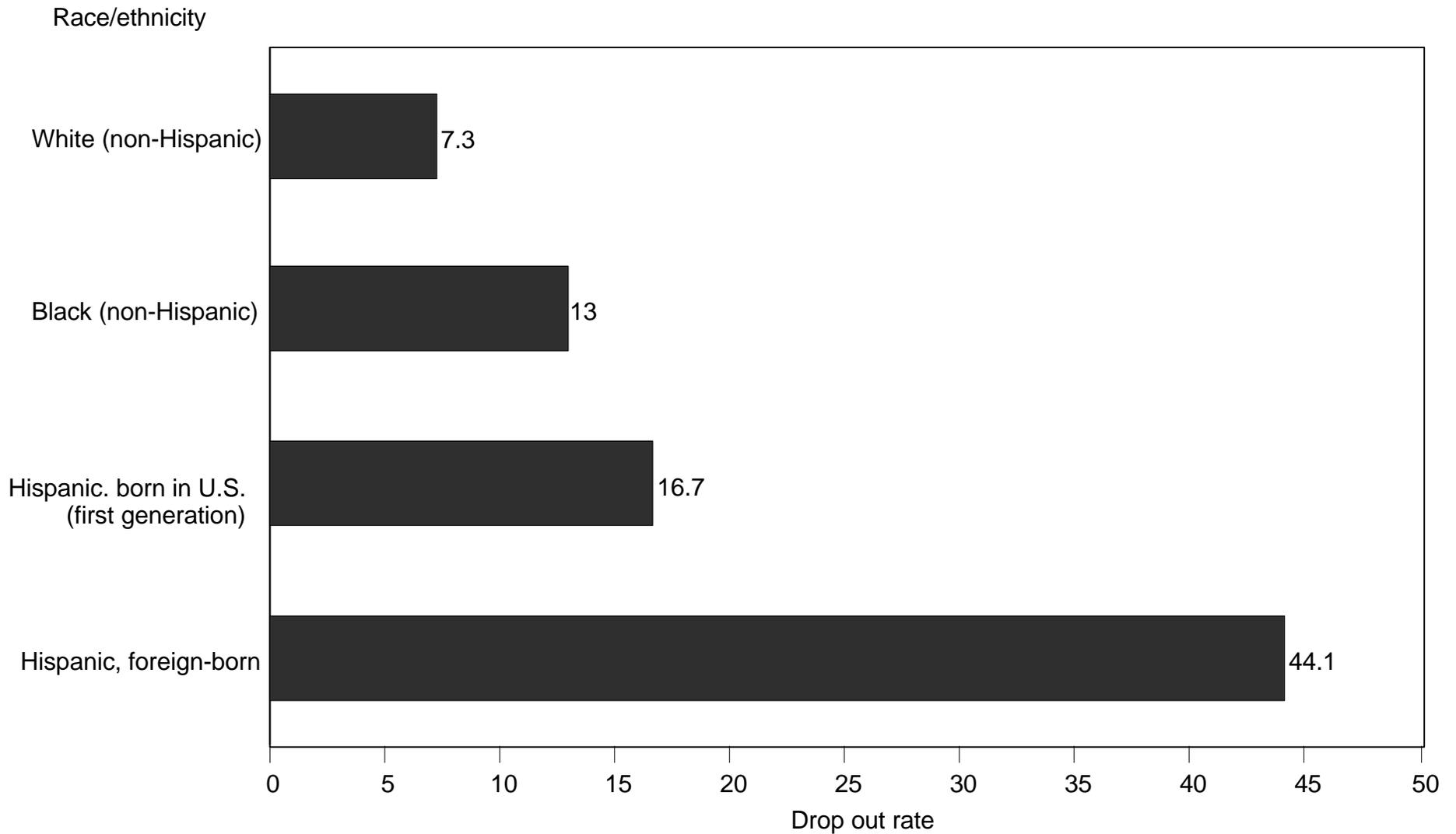
**Figure M. Percent of 16-to 24-year-olds Who are Dropouts, by Race/Ethnicity:
October 1972–96**



*The status dropout rate measures the proportion of individuals who are dropouts at any one given time, regardless of when they dropped out of school.

SOURCE: U.S. Department of Education, National Center for Education Statistics, *Dropout Rates in the United States: 1996* (based on the October Current Population Surveys).

Figure N. Percent of 16- to 24-year-old Dropouts, by Race/Ethnicity: October 1996



SOURCE: U.S. Department of Education, National Center for Education Statistics, *Dropout Rates in the United States: 1996* (based on the October Current Population Surveys).