Questions and Answers on 1996 NAEP Science Results
October 21, 1997

Background:
Today, the National Assessment Governing Board (NAGB) released a performance report on the 1996 National Assessment of Educational Progress (NAEP) in science, showing that less than 30% of U.S. students reach proficient levels of scientific knowledge and understanding, with 29% reaching the proficient level at grades 4 and 8, and 21% demonstrating proficiency at grade 12. The basic level -- representing partial mastery of fundamental knowledge and skills -- was reached by 67% of students in grade 4, 61% in grade 8, and 57% in grade 12.

NAEP is a representative sample test that gives information on how the U.S. and over 40 states are doing at the state and national levels, but provides no information for individual students and their families. The 1996 science NAEP was developed from a new content framework, so results cannot be directly compared to previous assessments. Achievement is reported based on levels for basic, proficient, and advanced performance approved by NAGB in August.

Q: Why don't U.S. students perform better on this test?

A: These tests are very challenging, but U.S. students *should* do better on them, and the results are a strong indication that we have a long way to go. States and school districts must continue to raise their expectations for student achievement in science, adopt rigorous curricula and materials, and make sure teachers get the preparation and training they need. Parents and schools must encourage students to lay a strong foundation for later achievement by taking the toughest math and science courses possible. The President's proposed voluntary national tests will help as well, focusing the attention of communities on helping students master the basics of reading and mathematics, both of which are abolutely essential for students to excel in science subjects.

Q: It appears that U.S. students do worse the longer they stay in our educational system. Is that accurate?

A: It's true that the proportion of students attaining at least the basic level is lower for 12th grade than it is for 4th grade: fewer 12th graders met NAGB's standards for 12th grade peformance than 4th grade students met standards for 4th grade performance. The bottom line is that we need to improve curriculum and instruction at all grades.

Q: How can it be that the U.S. is second in the world on the TIMSS 4th grade science test and yet such a large proporation of our students don't even reach the basic level?

A: These results are an indication that the NAEP science tests are genuinely challenging. While TIMSS shows that average U.S. performance is high relative to other countries, NAEP shows that there are still many students attaining low performance levels. If the U.S. is going to maintain or improve its position in future international comparisons,

more and more U.S. students will need to reach higher achievement levels on the rigorous U.S. NAEP science test.

Questions and Answers on 1996 NAEP Science Results
DRAFT -- October 21, 1997

Background: Today, the National Assessment Governing Board (NAGB) released a performance report on the 1996 National Assessment of Educational Progress (NAEP) in science, showing that less than 30% of U.S. students reach proficient levels of scientific knowledge and understanding, with 29% reaching the proficient level at grades 4 and 8, and 21% demonstrating proficiency at grade 12. The basic level -- representing partial mastery of fundamental knowledge and skills -- was reached by 67% of students in grade 4, 61% in grade 8, and 57% in grade 12.

NAEP is a representative sample test that gives information on how the U.S. and over 40 states are doing at the state and national levels, but provides no information for individual students and their families. The 1996 science NAEP was developed from a new content framework, so results cannot be directly compared to previous assessments. Achievement is reported based on levels for basic, proficient and advanced performance approved by NAGB in August.

Q: Why don't U.S. students perform better on this test?

A: These tests are very challenging, but U.S. students *should* do better on them, and the results are a strong indication that we have a long way to go. States and school districts must continue to raise their expectations for student achievement in science, adopt rigorous curricula and materials, and make sure teachers get the preparation and training they need. Parents and schools must encourage students to lay a strong foundation for later achievement by taking the toughest math and science courses possible. The President's proposed voluntary national tests will help as well, focusing the attention of communities on helping students master the basics of reading and mathematics, both of which are abolutely essential for students to excel in science subjects.

Q: It appears that U.S. students do worse the longer they stay in our educational system. Is that accurate?

A: It's true that the proportion of students attaining at least the basic level is lower for 12th grade than it is for 4th grade,but this doesn't mean that older students know less science. The achievement levels for the science assessments are set separately for each grade, and the tests have few questions in common. What the numbers do mean is that fewer 12th graders met NAGB's standards for 12th grade peformance than 4th grade students met standards for 4th grade performance. The bottom line is that we need to improve curriculum and instruction at all grades.

Q: How can it be that the U.S. is second in the world on the TIMSS 4th grade science test and yet such a large proporation of our students don't even reach the basic level?

A: These results are an indication that the NAEP science tests are genuinely challenging. While TIMSS shows that average U.S. performance is high relative to other countries, NAEP shows that there are still many students attaining low performance levels. If the U.S. is going to maintain, or improve its position in future intermational comparisons, more and more U.S.

students will need to reach higher achievement levels on the rigorous U.S. NAEP science test.

UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF THE SECRETARY

FAX TRANSMITTAL

TO: Mike Cohen / Bill Kincaid

ORGANIZATION: DRC

PHONE NUMBER:

FAX NUMBER:

FROM: Kevin Sullivan

PHONE NUMBER: 401-0831

MESSAGE: Science

CONFIDENTIALITY NOTICE

THIS TRANSMISSION IS INTENDED FOR AND RESTRICTED TO THE NAMED ADDRESSEE ONLY. IT MAY CONTAIN CONFIDENTIAL AND/OR PRIVILEGED INFORMATION. IF YOU RECEIVE THIS TRANSMISSION IN ERROR. YOU ARE NOTIFIED THAT YOU ARE PROHIBITED FROM READING, COPYING, OR DISSEMINATING THE TRANSMISSION. PLEASE CALL 202-401-3000 TO ARRANGE FOR RETURN OF ANY TRANSMISSION SET IN ERROR. THANK YOU.

13 **PAGE(S) TO FOLLOW**

THE NATION'S REPORT CARD

National Assessment Governing Board

National Assessment of Educational Progress

HOLD FOR RELEASE AT 9:30 a.m.
Tuesday, October 21, 1997

CONTACT Lawrence Feinberg
(202) 357-6942

LESS THAN 30 PERCENT OF U.S. STUDENTS REACH PROFICIENT LEVELS IN SCIENCE, NATIONAL ASSESSMENT BOARD REPORTS

Although about two-thirds of American students can demonstrate at least a basic level of scientific knowledge and understanding, less than 30 percent reach proficient levels of science achievement for their grade, the National Assessment Governing Board reported today.

The report, based on nationally-representative samples of students, indicates that just 21 percent can reach a proficient level of science achievement in grade 12, where standards are highest, compared to 29 percent at or above the proficient level for grades 4 and 8.

State-by-state data in terms of achievement levels are also included in the new report which describes the Board's standards for analyzing results of the 1996 science assessment of the National Assessment of Educational Progress (NAEP).

The report shows that very few students--3 percent at each of the three grades tested--have reached the advanced levels set by the Governing Board.

The independent, bipartisan 26-member Board is authorized by Congress to make policy for the National Assessment, the nation's only program that regularly measures the academic achievement of a cross-section of students in U.S. elementary and secondary schools. By statute, the Board is responsible for setting "appropriate student performance levels" for each grade and subject tested by NAEP.

Since 1991, similar achievement levels have been established by the Governing Board in four other subjects--mathematics, reading, U.S. history, and geography. A NAEP report on science was issued in May with scale-score comparisons only because the Board had not yet completed deliberations to adopt performance standards.

The proficient level is defined by the Board as representing "solid academic performance" which demonstrates "competency over challenging subject matter" for each grade tested.

-MORE-

800 North Capitol Street, N.W.
Suite 825, Mailstop 7583
Washington, D.C. 20002-4233
Phone: (202) 357-6938
Fax: (202) 357-6945

According to the new report:

o In state-level assessments, conducted in grade 8 only, the proportion of students at or above the proficient standard ranges from 41 percent in Maine, Montana, and North Dakota to 13 per cent in Louisiana, 12 percent in Mississippi, and 5 percent in the District of Columbia. Results are available for 40 states and four other participating jurisdictions.

o Nationwide, the basic level, showing partial mastery of fundamental knowledge and skills, was reached by 67 percent of students in grade 4, 61 percent in grade 8, and 57 percent in grade 12.

o Among the states, the proportion reaching basic, ranged from 78 percent in Maine and North Dakota to 39 percent in Mississippi and 19 percent in the District of Columbia.

Overall, about 130,000 students were tested in NAEP's science samples for 1996. For national results both public and private school students are included. The state comparisons are based on representative samples of public school students only. Separate results are reported for private schools in 18 states and Guam for which sufficient samples of students were tested.

"With achievement levels, the National Assessment is being transformed into a standards-based test," said Board Chairman Mark D. Musick, who also is president of the Southern Regional Education Board. "The Governing Board carefully defines how good is good enough. The assessment then reports on how many students reach these expectations."

The science assessment for 1996 was developed from a new content framework adopted by the Governing Board. Results cannot be compared to previous assessments. About 80 percent of student time was spent on open-ended items, the remainder on multiple-choice questions. Each test booklet included a hands-on science task or simple experiment.

"We are very aware that this NAEP assessment provides a new baseline for science achievement," Musick said. "It will be repeated in the year 2000 and four years after that. We hope it is a baseline that holds up."

The National Education Goals Panel--composed of governors, state legislators, Administration officials, and members of Congress--plans to use data from the new science achievement levels in its 1997 report, scheduled for release November 5.

Each achievement level has an expected score on the NAEP scale, a corresponding description of subject-matter knowledge and skills, and illustrative questions and answers. The science levels were approved by the Board in August after a year-long process that included panels of teachers, other educators, and members of the public from across the country. Several hundred persons and groups participated in the process, which was managed under contract by ACT (formerly American College Testing).

For both national and state results, the proportion of students at or above each achievement level is presented by gender, ethnic group, and parental education. Results also are given for low-income children who participate in Title I or free and reduced-price school lunch programs.

Patterns of performance across demographic groups are similar to those in other NAEP reports. However, state-level results show wide variations in how well students of similar background perform in different states. Even among the highest performing group--the children of college graduates--only about half the students in six states reach the proficient level.

In the national sample, of eighth grade students whose parents are college graduates, 39 percent reached the proficient level, compared to 33 percent for those whose parents had some education after high school, 18 percent for children of high school graduates without any post-secondary education, and 10 percent for the children of school drop-outs.

Overall, 37 percent of white eighth graders were at or above the proficient level, compared to 30 percent of Asians, 24 percent of American Indians, 11 percent of Hispanics, and 5 percent of blacks. Thirty-one percent of eighth grade boys reached the proficient level, compared to 27 percent of eighth grade girls.

The National Assessment Governing Board is an independent policy-making body composed of state, local, and federal officials, educators, and members of the public. The achievement levels committee is headed by Michael T. Nettles, Professor of Education and Public Policy at the University of Michigan and Director of the Frederick D. Patterson Research Institute of the United Negro College Fund.

In addition to the report, NAEP 1996 Science Performance Standards, the Board is distributing an eight-page tabloid summary of highlights for a general audience, called Science: What Do Students Know? Copies of both are posted on the NAGB web site on the Internet at http://www.nagb.org. They are also available free of charge from the Governing Board office, 800 North Capitol St., N.W., Washington, D.C. 20002, telephone (202) 357-6942.

# # #

NAEP 1996 Science Achievement Levels Facts

Title of report	*NAEP 1996 Science Performance Standards*
Grades assessed	National: Grades 4, 8, and 12 State: Grade 8
National sample sizes State sample size goals	Grade 4: 237 schools, 7,305 students Grade 8: 202 schools, 7,774 students Grade 8: 232 schools, 7,537 students Grade 8 only: 100 schools, 2,500 students per state
Jurisdictions assessed	47 jurisdictions, including 43 states, DC, Guam, and DOD schools (both domestic and overseas). Results were reported for jurisdictions that met the initial school participation rate of 70 percent (3 jurisdictions failed to meet this guideline).
Test formats	37% multiple choise, 49% short answer, 13% extended constructed-response. At least 1 hands-on performance task per student
Content areas measured	Life, physical, and earth sciences
Other scientific skills measured	Conceptual understanding, scientific investigation, and practical reasoning
How results are reported	Percent at or above basic, proficient, and advanced achievement levels
Descriptions of performance within achievement levels	What students know and can do
Background variables	gender, race/ethnicity, parental education, type of school, eligibility for Title I and free and reduced lunch programs
Time of data collection	National: October 1995 to May 1996 State: October to December 1995

National Assessment Governing Board

1996 NAEP Science Performance Standards

Data Presentation

Mary Lyn Bourque
Assistant Director for Psychometrics

Introduction

The National Assessment Governing Board is releasing today student performance data on the 1996 National Assessment of Educational Progress (NAEP) science assessment. These data are *similar* to those released by the National Center for Education Statistics (NCES) insofar as they are based on the 1996 science survey administered to a national sample of approximately 22,000 students in grades 4, 8, and 12, and to additional state or jurisdictional samples totaling 108,000 public and nonpublic school students in grade 8. They are *different* insofar as they report student performance against the achievement levels adopted by the Board, which interpret the NAEP data by focusing on the question, "How good is good enough."

The 1996 science assessment asked students to answer a sample of questions covering three major areas: physical, life, and earth science. Some questions were the conventional multiple choice type. However, more than 80% of the students' time was spent responding to open-ended questions. These questions required students to generate and write their answers, and in some cases, read graphs and charts and predict an outcome. Students spent a total of about 90 minutes answering the questions, and about one-third of that time they were required to engage in scientific investigations. These "hands-on tasks" as they are called, provided students an opportunity to work through a set of structured activities, using science equipment, that allowed them to make observations and draw interpretations and conclusions based on their work. These tasks engaged students in assessment activities which are more closely related to 'real science' than the traditional paper-and-pencil tasks..

Under the current legislation, the Board is mandated to develop student performance standards on the National Assessment through a national consensus process. Over the past 18 months, the Board has contracted with ACT (formerly known as American College Testing) to assist in developing a national consensus focusing on the question, "What *should* students know and be able to do in science?" Literally hundreds of teachers, administrators, curriculum specialists, scientists, and members of the public including parents and business leaders, have assisted the Board in developing the levels.

There are three levels for each grade assessed: *Proficient* (the central level, and the level to which all students should aspire) is defined as competency over challenging subject matter; *Basic* represents partial mastery; and *Advanced* signifies superior performance. These general definitions are further elaborated in terms of the kinds of knowledge and skills *in science* students should be able to know and do *at each grade level*. In addition, the levels are placed onto the NAEP scale (which for science is a 0 to 300 scale), so that the percentages of students

reaching the levels can be reported. Finally, there are some example questions for each level along with student responses which give the reader an idea of what is accepted as evidence of having reached each performance level.

The achievement levels in this report describe what student performance should be at three grade levels as measured by the National Assessment of Educational Progress. Today, the Board is also reporting to the American public on how successful students have been in reaching the student performance levels adopted by the Board. This report should be viewed as a companion report to the NCES report released in May, since it is based on the same data and identical analytical techniques.

National Results

Using the NAEP achievement levels as the yardstick, where do we stand as a nation in science? Nationally, 29% of the students in grades 4 and 8, and 21% in grade 12, were at or above the *Proficient* level. Of those, 3% of each grade also reached the *Advanced* level. Sixty-seven percent of grade 4 students, 61% of grade 8 students, and 57% of grade 12 students performed at or above the *Basic* level. The remaining percentages of students, 33% in grade 4, 39% in grade 8, and 43% in grade 12, were not successful in reaching any of the student performance standards on NAEP.

Gender Differences. In terms of gender differences, only grade 8 showed no difference in science performance between boys and girls. At grade 4, some difference was evident at the *Proficient* level, favoring boys. At grade 12, males outperformed females at all three achievement levels.

Racial/Ethnic Differences. NAEP reports on five major racial/ethnic subpopulations: Whites, Blacks, Hispanics, Asian/Pacific Islanders, and American Indians. These categories are based on students' self-reports of their background. At grade 4, the percentages of Whites, Asian/Pacific Islanders, and American Indians at or above *Basic* and *Proficient* were greater than the percentages for Blacks and Hispanics at or above these same levels. The sample sizes at the *Advanced* level were too small to make any statistical comparisons among these groups.

At the grade 8 *Basic* level, Whites, Asian/Pacific Islanders, and American Indian outperformed Blacks and Hispanics. At the *Proficient* level, Whites and Asian/Pacific Islanders outperformed Blacks and Hispanics, while Hispanics outperformed Blacks at the both the *Basic* and *Proficient* levels.

At the grade 12 *Basic* and *Proficient* levels, White and Asian/Pacific Islander students outperformed Black and Hispanic students. At the *Advanced* level White students outperformed Hispanic students.

Parents' Highest Education Level. Typically, students who reported their parents had attained higher levels of education did better than students whose parents had not completed such higher levels. For example, at both grades 4 and 8, a greater percentage of students whose parents had graduated college reached the *Advanced* level than those students reporting their parents had graduated high school. Similarly, at the *Basic* and *Proficient* levels, students who reported their parents had completed college or had some education after high school outperformed those who reported neither parent finished high school or that one parent had completed high school.

At grade 12, these differences were more numerous, with students whose parents had higher levels of education outperforming students whose parents had not completed higher levels. This same trend was evident at the *Basic* and *Proficient* levels among students who reported their parents had graduated college and those whose parents did not finish high school, had completed high school, or had some education beyond high school.

Type of School. NAEP reports on two major subpopulations with respect to the type of school students attend: public and nonpublic. In the latter category are included both Catholic and other nonpublic schools. At grade 4, students enrolled in nonpublic schools outperformed students in public schools at the *Basic* (82%/65%) and *Proficient* (42%/27%) levels, but not at the *Advanced* level. However, at grades 8 and 12, there were no significant differences between public and nonpublic school students' performance.

State Results

This report provides grade 8 student performance information for 44 jurisdictions, including 40 states, the District of Columbia, Guam, the Department of Defense domestic schools and the Department of Defense international schools. For purposes of making comparisons among the states and between the states and the nation, the state samples include *only* public school students, which can then be compared to the national public school sample.

The percentage of students reaching the *Proficient* level in the states ranged from a high of 41% in Maine, Montana and North Dakota to a low of 5% in the District of Columbia. In the national public school sample, 27% of the students performed at or above the *Proficient* level. Using the *Proficient* level as the comparison point, the accompanying map displays the 16 states/jurisdictions (coded red) in which the percentages of students at or above the *Proficient* level are significantly higher than that of the nation. Eleven states/jurisdictions did not perform differently from the national comparison group (coded green), while 17 states/jurisdictions performed significantly lower than the nation (coded orange). The remaining 10 jurisdictions (coded blue) either did not participate in the 1996 NAEP science assessment or their data are not being reported due to low participation rates.

Gender Differences. At the national level there were no significant differences between boys and girls at grade 8. However, there was more variability evident in the states. For example, in 12 jurisdictions boys outperformed girls at the *Proficient* level. In two states, Michigan and Utah, boys outperformed girls at all three achievement levels. And in five states, Michigan, Nebraska, New York, Utah, and Washington, boys outperformed girls at the *Advanced* level.

Racial/Ethnic Differences. The trends evident in the national data are apparent in the state data as well. In about 31 of the 44 jurisdictions Whites outperformed Blacks at the *Proficient* level. In 40 of the 44 jurisdictions Whites outperformed Hispanics at the *Proficient* level. The results are about the same for the *Basic* level as well.

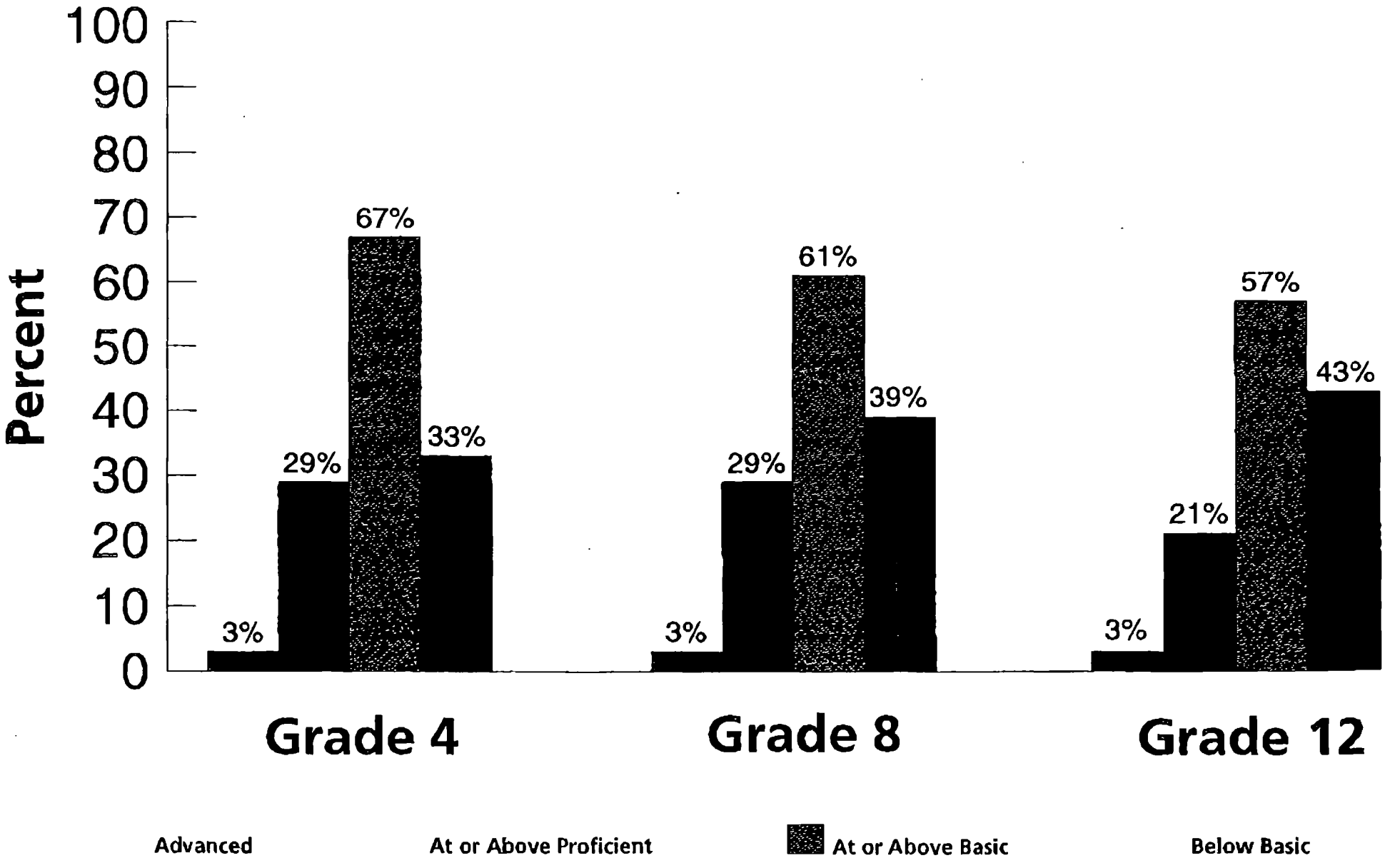
Parents' Highest Education Level. The trend in this area for the states is very similar to the results reported for the nation. Higher levels of parental education are generally associated with higher student performance. This is particularly true for student performance at or above the *Basic* and *Proficient* levels.

Type of School. Nineteen states/jurisdictions had a sufficient sample to report the data separately for public and nonpublic performance. Achievement differences at both the *Basic* and *Proficient* levels based on the type of schools students attend were evident in five states/jurisdictions, including Arkansas, Louisiana, Missouri, Montana, and Guam, where nonpublic school students outperformed public school students. At the *Basic* level only, seven states showed higher performance for nonpublic over public school students.

Student Performances

Finally, the report displays some items which are exemplars of the achievement levels adopted by the Board, as well as student answers to the open-ended questions. Three of these have been selected for display here. The first is a grade 4 *Proficient* exemplar, a multiple choice question. For this particular question, 76% of the students who were in the *Proficient* category correctly answered this question. The second question, a grade 8 *Advanced* exemplar, is a typical open-ended question. This is a good example of where students are given a data chart and are asked to interpret the data and make a prediction about what would happen to the temperature of the sand and water at about noon. Fifty-nine percent of the students in the *Advanced* level were likely to receive a score of 2 on that item. Finally, a grade 12 *Proficient* exemplar, where 68% of the students at the Proficient level are likely to get a score of 3. The question requires students to describe a test for telling the difference between salt water and lake water, and to explain how the test would work.

Percentage Attaining Science Achievement Levels for the Nation



Science Achievement Level Results for Grade 8 Public Schools

Percentages of Students At or Above Proficient Are Higher than the Nation

Percentages of Students At or Above Proficient Are Not Significantly Different from the Nation

Percentages of Students At or Above Proficient Are Lower than the Nation

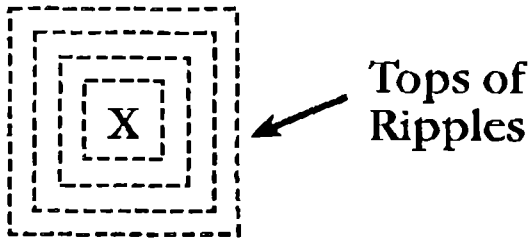
Did Not Participate

[1]Indicates that the state did not satisfy one or more of the guidelines for school participation rates

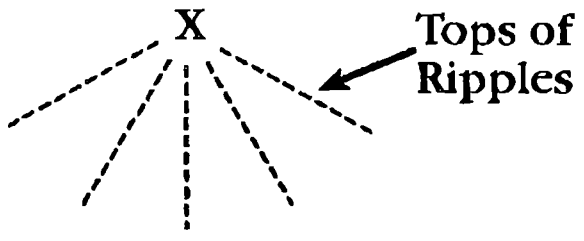
[2]Failed to meet initial school participation rate of 70 percent for public schools; public school results not reported

Grade 4 Proficient Exemplar

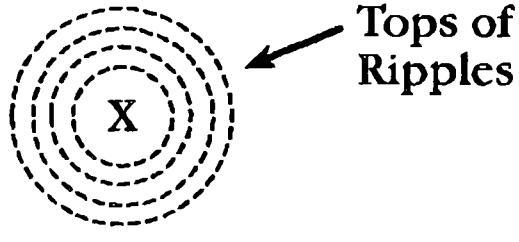
Ⓐ



Ⓑ



●



Ⓓ

You stand on the end of a boat dock and toss a small stone out into a pond of still water. Ripples form on the surface of the water. Which drawing shows what you will see when you look down at the water? (X marks where the stone enters the water.)

Percentage of students at the Proficient level answering correctly:

76%

Grade 8 Advanced Exemplar

Suppose that the experiment yielded the results shown in the table below.

Time (min)	0	1	2	3	4	5	6	7	8
Soil Temp (°C)	20	21	22.5	24	26	27.5	29.5	30.5	32
Water Temp (°C)	20	21.5	23	23.5	24	25.5	26	27.5	28.5

At a beach that has white sand, you measure the temperature of the sand and the temperature of the seawater at 9:00 a.m. You find that both have a temperature of 16°C. If it is clear and sunny all morning, what do the data from the experiment predict about the temperature of the white sand compared to the temperature of the seawater at noon?

The sand will become more warmer than the water by noon.

Explain your answer.

Well the sun is exposed to both the water and the sand. but the water will evaporate when it becomes heated and the sand will absorb the heat from the sun

Explain why the prediction based on the data might be wrong.

It might be wrong because the time data is incorrect.

SCORING GUIDE

4 = Lists and explains a prediction and explains why it may be wrong.

3 = Lists a prediction and either explains it or explains why it may be wrong.

2 = Lists a reasonable prediction.

1 = Provides the beginning of a prediction or explanation.

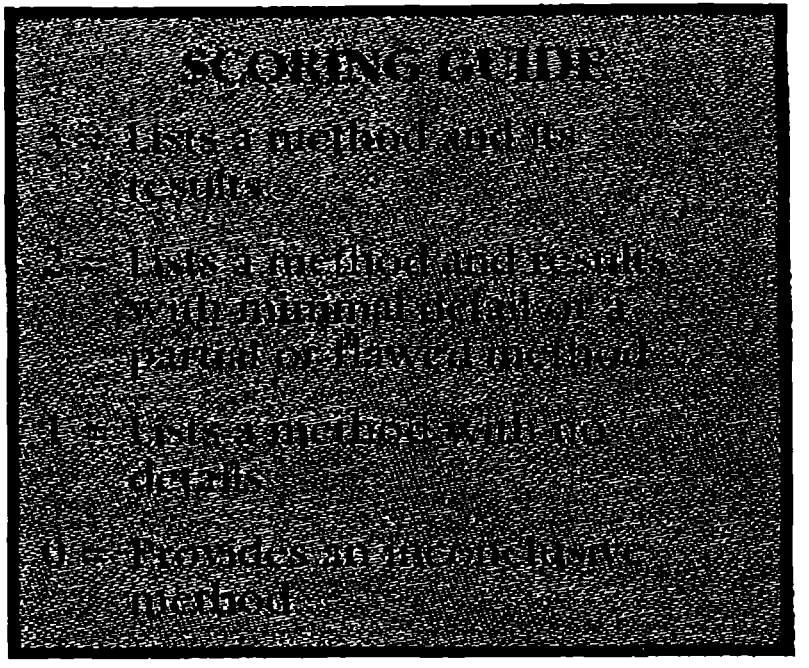
Percentage of students at the Advanced level receiving a score of 2:

59%

Grade 12 Proficient Exemplar

Some students were studying water in the environment. They filled one sample jar with ocean water and another sample jar with fresh water from the lake. The labels on the jars fell off and the water in both jars looked the same. Describe a test other than tasting or smelling the water that the students could do to determine which jar held the ocean water and which jar held the lake water. Explain how the test would work.

- the students could pour some water from each jar into a bowl and then place the bowl in the sunlight. After a few days they could come back and see which bowl had salt left in it after the water had evaporated and the one with the salt would be the one from the ocean.



Percentage of students at the Proficient level receiving a score of 3:

68%

Questions and Answers on 1996 NAEP Science Results
DRAFT -- October 20, 1997

Background: Today the National Assessment Governing Board (NAGB) released a performance report on the 1996 National Assessment of Educational Progress (NAEP) in science, showing that although about 2/3 of U.S. students can demonstrate at least a basic level of scientific knowledge and understanding, less than 30% reach proficient levels at grades 4, 8, and 12. At least the basic level -- representing partial mastery of fundamental knowledge and skills -- was reached by 67% of students in grade 4, 61% in grade 8, and 57% in grade 12. The 1996 science NAEP was developed from a new content framework, so results cannot be directly compared to previous assesments. Achievement is reported based on levels for basic, proficient and advanced performance approved by NAGB in August.

Q: Why don't U.S. students perform better on this test?

A: These tests are very challenging, but U.S. students *should* do better on them, and the results are a strong indication that we have a long way to go. States and school districts must continue to raise their expectations for student achievement in science, adopt rigorous curricula and materials, and make sure teachers get the preparation and training they need. Parents and schools must encourage students to lay a strong foundation for later achievement by taking the toughest math and science courses possible. The President's proposed voluntary national tests will help as well, focusing the attention of communities on helping students master the basics of reading and mathematics, both of which are abolutely essential for students to excel in science subjects.

Q: It appears that U.S. students do worse the longer they stay in our educational system. Is that accurate?

A: It's true that the proportion of students attaining at least the basic level is lower for 12th grade than it is for 4th grade,but this doesn't mean that older students know less science. The achievement levels for the science assessments are set separately for each grade, and the tests have few questions in common. What the numbers do mean is that fewer 12th graders met NAGB's standards for 12th grade peformance than 4th grade students met standards for 4th grade performance. The bottom line is that we need to improve curriculum and instruction at all grades.

Q: How can it be that the U.S. is second in the world on the TIMSS 4th grade science test and yet such a large proporation of our students don't even reach the basic level?

A: These results are an indication that the NAEP science tests are genuinely challenging. However, if the U.S. is going to maintain, or improve its position in future intermational comparisons, more and more U.S. students will need to reach higher achievement levels on the rigorous U.S. NAEP science test.

Q: At some grade levels (4 and 12) and achievement levels boys performed significantly better than girls. Do these findings make you more inclined to support experiments with single-sex schools or classrooms?

A: These findings certainly raise a concern that we may need to do more to insure that girls have equal opportunities to excel in science. However, we need more analysis to understand why these differences are showing up, and to clarify the kinds of steps that could be taken to address those differences.

Possible General Talking Points on Science Tests
Richard W. Riley, U.S. Secretary of Education
October 21st, 1997

- Science and math are increasingly important in this technology era and information age.

- This test of science is a challenging test including hands-on experiments and explaining answers in writing. It reflects the "real-world" use of science. The test and the achievement levels on this new science test demonstrate what we want our students to achieve. For many students they have a distance to go, but now we have a benchmark from which to make progress.

- Yesterday, our Department released a study that showed that taking algebra and geometry early in the secondary school experience was very important in opening the gate to other important high school courses to prepare for college, such as more advanced science and math courses. Those findings give us guidance also here today to boost science achievement in high schools -- where we must do better in science.

- The performance of 12th grade students in science appears to reflect the fact that not enough students take chemistry and physics and as result, they won't do as well on the physical science questions and concepts on the test. It is essential to begin on the road to physics and chemistry by encouraging students to take algebra in eight grade and build on it with more advanced science and math courses all through high school.

- So much of science in the 8th grade and 12th grade also depends on a solid foundation in the basics of reading and math. That is why the President's proposal for a voluntary national tests for 4th grade reading and 8th grade math is so important to boost achievement overall in the long term in a whole host of subjects - including science in both the 8th grade and 12th grade where we appear weaker.

A. Kolstad October 20, 1997 (1:24PM)

Q's and A's for the Secretary of Education and the Commissioner of Education Statistics

1. *Why is the science assessment important?*
Scientific literacy is essential for successful functioning in today's society and is a key to the continued leadership of our nation in the world. This assessment focuses on the science content that we believe students in this nation should know, and on the specific skills that our students must master to have an understanding of science. Last November, we reported that our nation's 8th graders placed above the international average in their knowledge of science, compared with about 45 other countries, though students in several countries were ahead of the United States in their level of science education (based on the results of the Third International Mathematics and Science Study).

2. *According to this assessment, how well are we doing in Science?*
The data in the *NAEP 1996 Science Report Card* released in May did not give us a yardstick to tell us how well we are doing, nor were trend comparisons possible. This report for the first time provides that yardstick in the form of achievement levels, which are intended to help answer the question of "How good is good enough?" The National Assessment Governing Board believes that all students should reach the *Proficient* level, and that the *Basic* level represents only partial mastery and should not be the desired goal.

   a. *How did the nation's students perform?*
   Twenty-nine percent of 4th and 8th grade students and 21 percent of 12th grade public and nonpublic school students performed at or above the *Proficient* level, defined as demonstrating their competency over challenging subject matter. This includes 3 percent of students at each grade level who met the *Advanced* standard for superior performance. This is not good enough. American students should do better in science.

   b. *When will it be possible to track trends in science performance according to this way of assessing science?*
   The next science assessment is scheduled for the year 2000, and future assessments are planned on a 4-year cycle. The next NAEP trend assessment, which has tracked science performance since 1969 according to a less comprehensive definition of science, will be conducted again in 1999.

   c. *How did students perform in the states?*
   - Sixteen jurisdictions (ME, MT, ND, WI, MA, MN, CT, IA, NE, VT, WY, CO, MI, OR, UT, and DoDDS) had greater percentages of public school students than the nation who were at or above the proficient level. Maine, Montana, and North Dakota tied for having the highest percentage of students (41 percent) achieving at or above the proficient level.
   - Seventeen jurisdictions (AZ, KY, AK, TN, DE, FL, GA, WV, CA, NM, AL, SC, HI, LA, MS, Guam and DC) had smaller percentages of public school students than the nation who were at or above the proficient level. The lowest

percentages were in Mississippi, with 12 percent reaching the proficient level or better, Guam at 7 percent, and the District of Columbia at 5 percent.

- Eleven jurisdictions (AK, IN, MO, NY, VA, WA, RI, MD, NC, TX and DDESS) had percentages of public school students at or above the proficient level that were not statistically different from that of the nation. The national percentage of public school students at or above the proficient level was 27 percent.

d. *Why do you think the states that had fewer students at or above the proficient level were mostly clustered in the southeast?*
This pattern of lower performance in the southeast has also been observed in previous NAEP assessments in different subjects. The southeast has more poverty and a history of lower educational attainment than other regions, and there is an association between poverty and academic performance.
However, I think that the important thing to concentrate on is improvement. While we cannot compare these states' performance in science with that for previous years, we noted with our mathematics assessment that several southeastern states had improved since the previous assessment in 1992. In fact, North Carolina was the state that improved the most in 8th grade.

3. *Do older students know less science?*
While it is true that the proportion of students below the basic level increases from fourth to eighth to twelfth grades, this does not mean that older students know less science. There were not enough common questions in science to put the three grades on a common scale, so the scale scores of the three grades are not comparable. In addition, the achievement level standards are set separately for each grade. What those numbers mean is that fewer twelfth grade students met our standards for twelfth grade performance than fourth grade students met our standards for fourth grade performance.

4. *How is this report different from the NAEP 1996 Science Report Card issued in May?*
The *NAEP 1996 Science Performance Standards* report differs from the May report in that this report balances the findings against judgmental standards for achievement. The science assessment results were issued in two companion reports rather than an integrated report because the achievement levels were not ready in time to be included in the initial publication of the *NAEP 1996 Science Report Card*. This came about because the Board found the levels resulting from its achievement-level setting process did not meet its test of reasonableness. Developing and implementing an alternative approach to setting achievment levels took time.

5. *How were the achievement levels set for the science assessment?*
In the final analysis, the Board exercised its judgement about where the achievement levels should be set on the NAEP scales to satisfy the reasonableness criterion. In deciding on where to set its levels, the Board took into account other relevant information, including achievement levels already set in other NAEP subjects, 1996 Advanced Placement (AP) results for twelfth grade students, and information about

eighth grade students gathered from the Third International Mathematics and Science Study.

6. *Why do the achievement level descriptions now say that students at each level are "able to do" certain things? Didn't previous NAEP reports say that the students at given levels "should be able to" do certain things? Where does the "should" come in?*
The Board believes that all students should perform at the *Proficient* level. The descriptions of what students know and can do within each achievement level is based on an anchoring process that NAEP had used in reporting performance prior to the adoption of achievement levels. The difference is that scale anchoring in NAEP used to be done with respect to arbitrary points along the scale, whereas in this assessment, the anchoring was done with respect to the cutpoints selected by the Board's judgment.

7. *Does NAGB plan to use the new achievement-level process in the future, or only when the normal process produces unreasonable results?*
The answer is not yet clear. NAGB expects to review its procedures for setting achievement levels and will decide later how to proceed.

8. *When will the achievement levels stop being used on a developmental basis?*
The law requires that when the Commissioner of Education Statistics determines that the achievement levels are reasonable, valid, and informative to the public the achievement levels can stop being used on a developmental basis. Until the Commissioner make that determination, the law requires all NAEP reports to make clear their developmental status. This report does state that the achievement levels are being used on a developmental basis. The Commissioner will be aided in his determination by the report of a National Academy of Science evaluation of NAEP that is currently being conducted and is scheduled to report in late 1998.

9. *How expensive was it to incorporate the hands-on activities, and was it worth the expense?*
In the 1996 science assessment, about 130,000 students each conducted a hands-on activity. The cost of creating and providing the materials, as well as scoring the results, was about 98 cents per student, or $1.25 million overall. It was worth the expense to do this because we feel that a very important part of assessing students' abilities in science is to find out whether they can actually do science. Providing these hands-on activities gave students a chance to demonstrate that they could work with the materials of science, make observations, and draw conclusions from them.
In future assessments, the redesigned NAEP will continue to use and report on hands-on tasks. However, some savings can be obtained by not requiring every student to respond to hands-on tasks and using the same kind of matrix sampling as that used for other types of tasks.