ETS POLICY NOTES

News from the ETS Policy Information Center

Volume 11, Number 1 Educational Testing Service Fall 2002

A Deeper Look at NAEP Science Results

Introduction

Persistent and disturbing differences in academic achievement among U.S. students, usually apparent when students are grouped by race/ethnicity and/or some measure of socioeconomic status, have received renewed attention from education policymakers. This focus on the achievement gap in the U.S. education system carries forward Goal 3 of the national education goals that students should demonstrate knowledge in "challenging subject matter," with the objective that "the academic performance of all students... will increase significantly in every quartile, and the distribution of minority students in every quartile will more closely reflect the student population as a whole." This goal is renewed in the No Child Left Behind Act of 2001, which focuses on raising the academic achievement of all students, setting a minimum threshold of performance toward which the lowest achieving subgroups of students must show progress on assessments in use in each state.

The major tool available to measure the overall achievement of U.S. students is the National Assessment of Educational Progress (NAEP), which regularly reports average scale scores and the percentage of students reaching specific "achievement levels" set by the National Assessment Governing Board.[1] National

[1] For more information on NAEP, go to http://nces.ed.gov/nationsreportcard.

This Issue — A Deeper Look at NAEP Science Results

This issue of *ETS Policy Notes* presents the results of a special analysis of NAEP state-by-state science data from 1996 and 2000 carried out by ETS under contract with the National Center for Education Statistics. The focus is on changes in the gap in scores between different groups of students.

Readers interested in more general results from the 2000 NAEP science assessment should see *Science Highlights 2000*, available from the National Center for Education Statistics (http://nces.ed.gov/nationsreportcard/science/results/index.asp).

Editor's note: This issue of *ETS Policy Notes* was written by Paul Barton, with data analysis provided by David Freund and Mei-jung Lin.

and state results in science were reported in November 2001.[2]

This issue of *ETS Policy Notes* provides a more in-depth examination of the gaps in science achievement than is provided in regular NAEP reports. It is based on a special analysis of state-by-state NAEP public school data from the 1996 and 2000 science assessments for Grade 8 conducted by ETS under contract with the National Center for Education Statistics. Thirty-three states participated in both assessments. Examining these state data can reveal patterns of performance that can be masked in national analyses. In addition to looking at changes in overall average scores and in the percentage of students reaching the "proficient level,"[3] this issue of *ETS Policy Notes* describes:

- Changes in achievement among top scorers (fourth quartile) between 1996 and 2000
- Changes in achievement among low scorers (first quartile)
- Changes in the gap between the top and bottom quartiles
- Changes in the gap between White and minority students[4]
- Changes in the gap between poor and nonpoor students[5]

A Deeper Look at the Data

While there were a few bright spots, the results over this four-year period of national concentration on standards-based reform were disappointing.

Figure 1 shows the change in performance for each state, on each of the seven indicators. Table 1 summarizes these results.[6]

Table 1 reveals that:

- There were **no** improvements on three indicators — the gap between the top and bottom quartile, the gap between White and minority students, and the gap between poor and nonpoor students.
- The gap **grew** in 11 states between the top and bottom quartile and in seven states between poor and nonpoor students.

[2] http://nces.ed.gov/nationsreportcard/pdf/main2000/2002452.pdf.

[3] NAEP defines this level as representing solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

[4] Minority students in this analysis are Black and Hispanic.

[5] While NAEP does not include a poverty measure, it does report the percentage of students who are eligible for a free or reduced-price lunch. This measure is used as a proxy for "poor" students.

[6] All differences noted are statistically significant based on *t*-tests.

Figure 1: Changes in NAEP Eighth-Grade Science Scores, 1996 to 2000

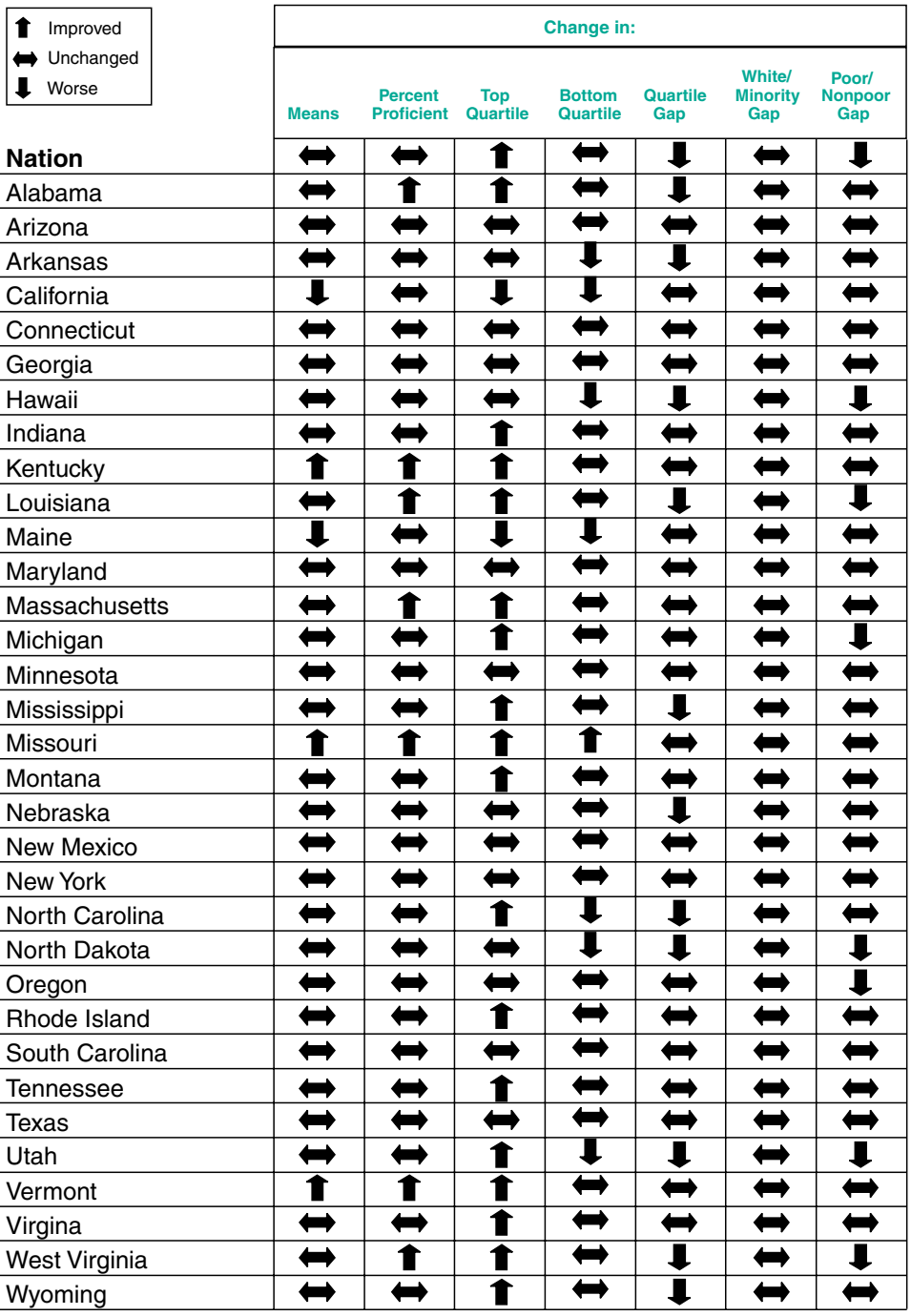


Table 1: Change Between 1996 and 2000 in NAEP Eighth-Grade Science Achievement on Seven Indicators

Change, 1996 to 2000 in:	States Improving	States Unchanged	States Worse
Average Score	3	28	2
Percent "Proficient"	7	26	0
Top Quartile	17	14	2
Bottom Quartile	1	25	7
Gap - Top and Bottom Quartile	0	22	11
Gap - White and Minority	0	33	0
Gap - Poor and Nonpoor	0	26	7

- Only one state showed an improvement for the **bottom** quartile of students.
- However, 17 states registered an improvement for the **top** quartile of the student population, and
- Seven states increased the percentage of students reaching the "**proficient**" level, although only three states raised the statewide average scores.

Table 2 provides another summary of how the states fared on the seven indicators of change that were used in this analysis.

As shown, none of the states improved on all seven or six or five of the indicators. Just one state improved on four of the seven indicators, and just two states improved on three indicators. Four states improved on two of the indicators and 10 on a single indicator. While no state did worse on four or more of the six indicators, six states did worse on one indicator, four states did worse on two indicators, and five states did worse on three indicators.

These results underscore the importance of examining multiple indicators of state performance when analyzing NAEP data. A look just at the percentage reaching the "proficient" level of performance would fail to disclose that no progress was made among those students in the bottom quartile of the score distribution. Without looking at the quartiles separately, we would not know the significant improvements in average scores for the top quartile of the students. And as we can see from Figure 1, the gap between the top and bottom quartiles increased in 11 states. In two of these states, scores dropped for the bottom quartile and rose for the top quartile, and in three states, scores declined in the bottom quartile without changing in the top quartile.

Table 2: State Performance on Seven Indicators of NAEP Science Achievement, 1996 - 2000

Number of Indicators	States Improving	States Unchanged	States Worse
1	10	0	6
2	4	0	4
3	2	4	5
4	1	8	0
5	0	5	0
6	0	7	0
7	0	9	0

The overriding concern is that there was little improvement in science achievement; there was no change in the nation as a whole in average scores or the percent reaching proficient—although a few states managed to pull average scores up.[7] There was an increase in scores in the top quartile for the country as a whole and in the top quartile in 17 states.

What is also noteworthy is that something in the dynamics of our educational system during this period was resulting in improvement in the scores of the best students, and deterioration—or no change—in the scores of the lowest scoring students, resulting in an enlargement of the score gap between the top and the bottom quartiles. In an earlier report on fourth grade reading (1992-1998), there was a similar result; scores declined in the bottom quartile in 18 states, while they rose in the top quartile in 12 states, with the gap between the top and bottom enlarging in 16 states. This was not the case for mathematics, where there were generally gains at both the top and bottom.[8]

There is no basis for speculation from these data as to what caused these growing gaps. However, the tracking of state-by-state progress points out the need to delve more deeply into the gaps than just examining statewide averages

[7] When nonpublic schools are included, there was a small but statistically significant increase in the percentage of eighth-graders reaching the "proficient" level or higher between 1996 and 2000.

[8] Paul E. Barton, *Raising Achievement and Reducing Gaps: Reporting Progress Toward Goals for Academic Achievement*, A Report to the National Education Goals Panel, March 2001. Available from http://www.ets.org/research/pic/raising.pdf

or the percentage reaching a particular score cut-point.

Persistent and Large Gaps

The previous section has summarized the situation in regard to whether the gaps are narrowing or widening over this four-year period. Figure 2 shows the size of these gaps for the participating states, for three indicators, in 2000. Typically, the gap between poor and nonpoor students is the smallest, and the gap between White and minority students is a little larger. However, the gap between the bottom and top quartile of all students is much larger, about 2 1/2 times the size of the gap between White and minority students. For the 2000 national results, these gaps were 32, 37, and 88 points, respectively. These are large gaps, and they have proven to be unyielding.

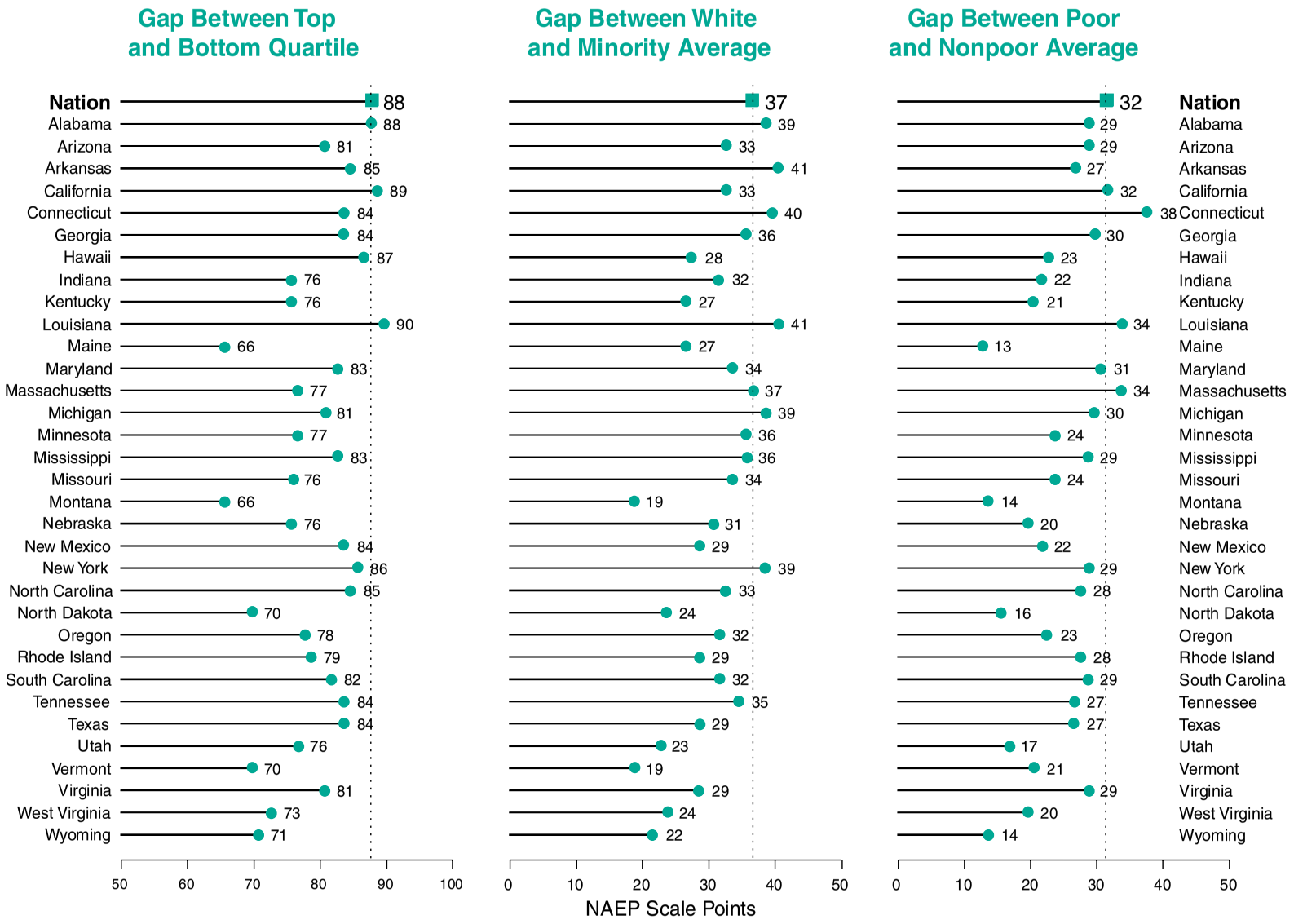
These gaps vary, state-by-state, as can be seen in Figure 2. However, no significance should be given to relatively small differences among states, given the errors due to the sampling approach in the NAEP assessment.

* * * * * * *

NAEP provides a rich source of information about elementary and secondary education for the nation and for the participating states. Measuring what school students know and are able to do is a little like wanting, from outside, to know what is inside a very large house. You look in one window and see a little, and you look in another window and see a little more. But to know what is inside the house, you have to look in all the windows. This report looks in several windows, but by no means all of them.

Another important dimension is the measurement of growth in achievement from the fourth to the eighth grade; in other words, how much students raised their scores over the four-year period. This is possible in math and reading where NAEP has been operating longer at the state level. This approach is used in *Growth in School: Achievement Gains from the Fourth to the Eighth Grade*, by Paul Barton and Richard Coley, a publication of the ETS Policy Information Center (ftp://ftp.ets.org/pub/res/growsch.pdf). In terms of state accountability testing, only Tennessee measures growth in knowledge during a year of schooling; the other states compare students at the end of a school year with their counterparts in previous years.

Figure 2: Gaps in Eighth-Grade NAEP Science Scores, 2000



Recently Published by the ETS Policy Information Center…

www.ets.org/research/pic

- *The Closing of the Education Frontier?*
- *Meeting the Need for Scientists, Engineers, and an Educated Citizenry in a Technological Society*
- *An Uneven Start: Indicators of Inequality in School Readiness*
- *Staying on Course in Education Reform*
- *Facing the Hard Facts in Education Reform*
- *The Twin Challenges of Mediocrity and Inequality: Literacy in the U.S. from an International Perspective*

ETS Policy Notes is published by the
ETS Policy Information Center
Educational Testing Service
Rosedale Road, 04-R
Princeton, NJ 08541-0001
(609) 734-5694

e-mail: pic@ets.org

www.ets.org/research/pic

Editor: Richard J. Coley

88503-02536 • Y102E6 • Printed in the U.S.A.
I.N. 995824