4 Comparing Changes in Achievement

A central concern to both state and NAEP assessment programs is an examination of achievement trends over time (e.g., USDE 2002, NAGB 2001). The extent to which NAEP measures of achievement trends match states' measures of achievement trends may be of interest to state assessment programs, the federal government, and the public in general. The purpose of this section is to make direct comparisons between NAEP and state assessment changes over time.

Unlike state assessments, NAEP is not administered every year, and NAEP is only administered to a sample of students in a sample of schools in each state. NAEP sample schools also vary from year to year. For this comparison report, our comparison of changes in mathematics achievement is limited to those between the 1999-2000 and 2002-2003 school years (i.e., between 2000 and 2003). For research purposes, analysts may wish to examine trends in earlier NAEP years (e.g., 1991-1992 and 1995-1996), but the NLSLSASD does not have sufficient state assessment data for those early years to warrant inclusion in this report.

To make meaningful comparisons of gains between NAEP and the state assessments, we included only the NAEP sample schools for which state assessment scores were available in this trend analysis.[27] This allows us to eliminate effects of random or systematic variation between schools in comparing NAEP and state assessments.

There are many states for which we did not have scores in multiple years and so could not measure achievement changes over time. In addition to these states, there are others for which we could not use 2003 scores for the trend analysis because they changed their state assessments and/or primary standards in 2003; changes in percentages meeting the primary standards from 2000 to 2003 will not reflect their actual changes in achievement. Therefore, these states are excluded from the analysis; they are listed below along with the reasons for exclusion:

27. These schools were weighted, according to NAEP weights, to represent the state. The assumption being made by using these unadjusted NAEP weights is that any NAEP school without state scores would have state scores averaging close to the state mean. That can be tested by comparing the NAEP means for the schools with and without state scores in each state. It is our belief that there is so few of these schools that it doesn't matter since we are matching close to 100 percent of the schools. Moreover, the analyses such as standard estimation are based on the subset of schools with both scores, so they are not biased by the omission of a few NAEP schools.

- California: changed assessment in 2003
- Indiana: changed assessment in 2003
- Maryland: changed assessment in 2003
- Michigan: changed assessment in 2003
- Nevada: changed assessment in 2003
- Texas: changed assessment in 2003
- Virginia: changes performance standards every year and no longitudinal equating from year to year done.
- Wisconsin: set new performance standards in 2003

It is important to note that changes in percentage meeting the primary standards may be affected by ceiling effects. In other words, if a state sets a relatively low standard and many schools in the state show very high percentages of students already meeting the standard in the base year, there will be little "room to grow" for these schools. The state would be less likely to show positive achievement trends, not because students are not learning, but because many students have already met the standard in the base year.

Finally, all significance tests are of differences between NAEP and state assessment results. The comparisons between NAEP and state assessment results in each state are based on that state's primary state standard. This means that the standard at which the comparison is made is different in each state. For this reason, comparisons between states are not appropriate.

In table 7, we summarize the average of, and variation in, changes in achievement over time on NAEP and the state assessments across states in terms of percentages achieving the state primary standards.[28] The state primary standard is, in most cases, the standard used for reporting adequate yearly progress, in compliance with *No Child Left Behind*. We *rescored* NAEP in terms of percentages meeting the state primary standards, because comparisons of trends at differing locations in the distribution of achievement are not easily interpretable. To rescore NAEP for trend comparisons, we estimated the location of the state primary standard on the NAEP scale in the initial trend year. In that year, the NAEP and the state assessment percentages match by definition. Finally, the State Profile section of this report (Appendix D) compares trends for multiple state standards, not only the single primary standard.

28. For Alabama, Tennessee, and Utah, we did not have percentages meeting a state standard; instead, for these states we report trends in percentile ranks.

Table 7. Mathematics achievement gains in percentage meeting the primary state standard in grades 4 and 8: 2000 and 2003

	Grade 4 2000 to 2003 (24 states)		Grade 8 2000 to 2003 (22 states)	
Statistic	**State**	**NAEP**	**State**	**NAEP**
Average gain	7.0	9.9 *	3.0	3.6
standard error	1.89	2.40	1.43	2.06
Between-state standard deviation	9.12	5.27 *	4.98	3.21 *
standard error	0.38	0.45	0.37	0.44

* NAEP gains are significantly different from gains reported by the state assessment ($p<.05$).

NOTE: Primary standard is the state's standard for *proficient* performance. State assessment gains are recorded here for the schools that participated in NAEP. Gains are weighted to represent the population in each state. Averages are based on states with scores on the same tests in the two years.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 and 2003 Mathematics Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

Averaged over the states in which gains from 2000 to 2003 could be compared, mathematics achievement gains reported by NAEP are larger than those in state assessment in both grades 4 and 8, although the differences are not statistically significant at grade 8. These results are reversed in reading: that is, reading achievement gains reported by state assessments are larger than those in NAEP in both grades 4 and 8 (McLaughlin et al., 2008).

There are many possible explanations for the differences in gains measured by NAEP and state assessments from 2000 to 2003 that are associated with differences in testing. We have constructed the comparisons to remove two important sources of error, by comparing trends for the same sets of schools and at the same standard level. Other factors to be considered include (1) differences in changes in accommodations provided on the two assessments; (2) differences in changes in the student populations in each sampled school (e.g., NAEP vs. all fourth graders); (3) differences in changes in motivation (low stakes/high stakes); (4) differences in test modality (e.g., multiple choice vs. constructed response); (5) differences in time of year; and (6) a recalibration of the state assessment between trend years (of which we are not aware).

In addition to differences in gains measured by NAEP and state assessments, variations in gains across states are also of interest. As shown in table 7, the gains in percentages meeting the primary standards measured by state assessment vary substantially between states. The standard deviations of these gains vary from five to nine percentage points. However, these differences may overestimate the actual variation in gains in different states. The standard deviation of gains between states is smaller when they are measured by a common assessment, NAEP.

Interpreting these variations requires caution. Gains are measured at different points on the achievement continuum in different states; therefore, the gains are not comparable across states. However, we believe that the search for common trends

across states provides us with valuable information to make descriptive statements about general patterns in the nation.

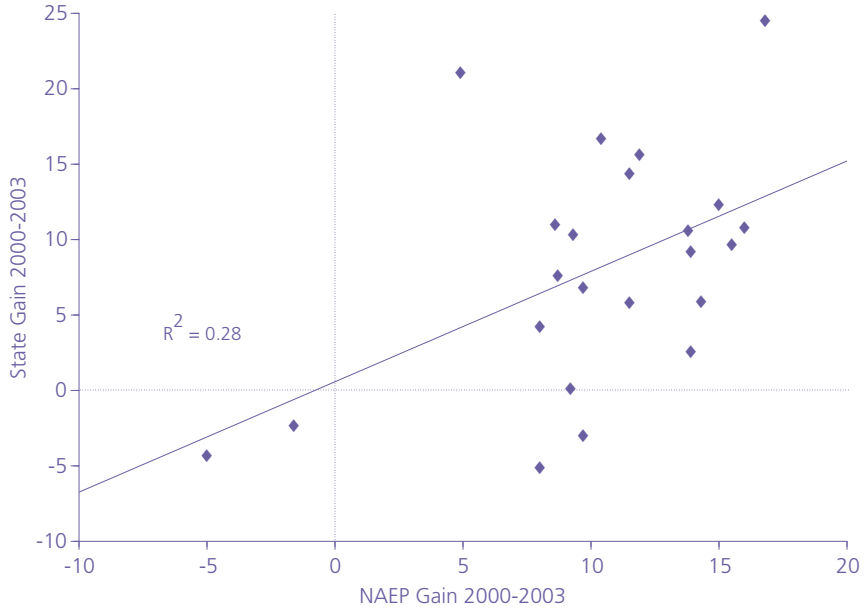
It is possible that the greater variation in state assessment gains than in NAEP gains is due to different states' measurement of unique aspects of mathematics achievement not fully addressed by NAEP. However, an alternative hypothesis must be considered before searching for the unique aspects of mathematics achievement measured in states with relatively large gains: that is, a substantial portion of the variation in state assessment gains may be due to methodological differences in the way that state assessments measure gains.

If the NAEP and state assessment gains are correlated positively, both assessments are likely to identify the same schools as making achievement gains. In other words, a school identified as increasing achievement by the state assessment is likely to be identified as increasing achievement by NAEP as well. To investigate whether the NAEP gains and state assessment gains are related, we present below scatter plots between the NAEP and state assessment gains for both grades 4 and 8.

Figure 17 indicates that the relationship between the NAEP and state assessment gains in grade 4 mathematics is positive, with R^2 of .28 and the correlation coefficient of .52 (with *p*-value<.05). This means that, in grade 4 mathematics, states that increased the percentage meeting the primary standard in the state assessment tend to be the states that increased the corresponding percentage in NAEP. Grade 4 mathematics is the only grade and subject that showed a relationship between the NAEP and state assessment gains: the results for grade 8, displayed in figure 18, and the reading results (McLaughlin et al. 2008) suggest that states in which the percentage of students meeting the primary reading standard based on the state assessment increased are not necessarily the states in which the percentage meeting the primary standard on NAEP increased.

Since the state assessment results vary significantly from state to state, it is important to identify particular states where the NAEP and the state mathematics assessment trends differ. A search for explanations of the different results must begin with identification of states in which they differ. State-by-state comparisons between NAEP and state assessment measurements of mathematics achievement trends are presented in the following tables.

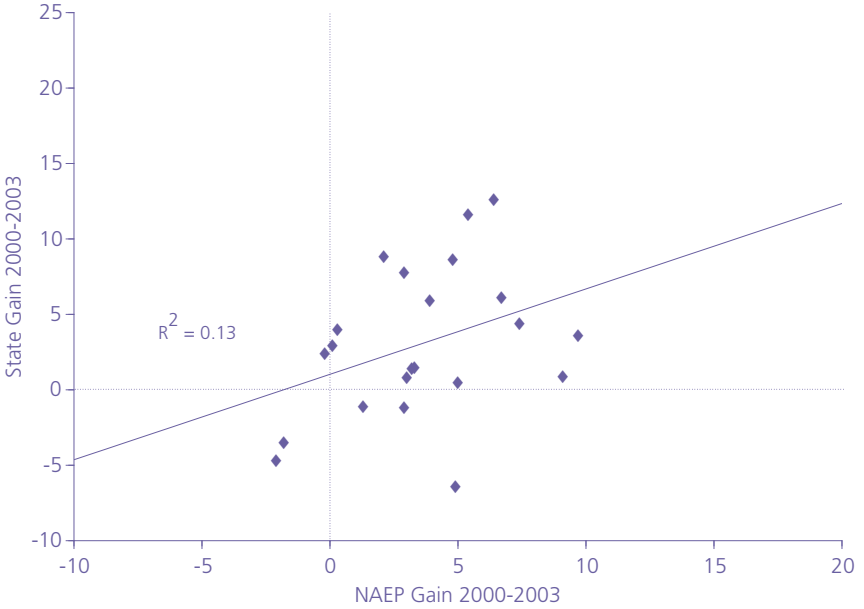
Figure 17. Relationship between NAEP and state assessment gains in percentage meeting the primary state grade 4 mathematics standard: 2000 to 2003



NOTE: Primary standard is the state's standard for *proficient* performance.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 and 2003 Mathematics Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

Figure 18. Relationship between NAEP and state assessment gains in percentage meeting the primary state grade 8 mathematics standard: 2000 to 2003



NOTE: Primary standard is the state's standard for *proficient* performance.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 and 2003 Mathematics Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

Table 8 summarizes state-by-state trends for grade 4 mathematics achievement. Eleven out of 24 states increased percentages meeting the primary standard in NAEP significantly more than in their state assessments from 2000 to 2003; these are Arizona, Connecticut, Hawaii, Louisiana, Maine, Massachusetts, Missouri, Oklahoma, South Carolina, Vermont, and Wyoming. On the other hand, in Arkansas, Ohio, and Rhode Island, state assessments found significantly greater gains from 2000 to 2003 in mathematics than NAEP did.

Table 9 summarizes state-by-state trends for grade 8 mathematics scores. Between 2000 and 2003, NAEP found significantly larger mathematics achievement gains than their state assessments in six out of 22 states; these are Connecticut, Hawaii, Massachusetts, Missouri, South Carolina, and Wyoming. All six states observed the same trends in their grade 4 mathematics results: that is, significantly greater gains in NAEP than in their state assessments. On the other hand, in Arkansas, Georgia, Illinois, New York, and Rhode Island, the state assessments measured significantly greater gains from 2000 to 2003 in the percentage meeting the state primary standard than NAEP did.

Do the states in which we found significant discrepancies between NAEP and state assessment trends differ from the other states? One reasonable explanation for trend differences is that the state assessments do not identify the same schools as high and low achieving that NAEP does. Since we *rescored* NAEP in terms of percentages meeting the state primary standards so that the NAEP and the state assessment percentages match in the base year (i.e., 2000), we address this issue by comparing a) the correlation between 2003 NAEP and state assessment results (table 4) in the states with significant discrepancies (in tables 8 and 9) to b) the correlation in other states without such discrepancies.

Overall, there is no noticeable difference between the states with significant discrepancies and those without such discrepancies.[29] In the grade 4 results, the average correlation between NAEP and state assessments in 2003 is .72 for the states with significant discrepancies in gains and .71 for other states without such discrepancies. Patterns are very similar in grade 8. The average correlation between NAEP and state assessments is .82 for the states with significant discrepancies in gains from 2000 to 2003 and .74 for other states without discrepancies. These results indicate that in both grades 4 and 8, there is no relationship between: a) the tendency for the two assessments to identify the same schools as low achieving and high achieving in 2003, and b) the sizes of discrepancies in gains as measured by NAEP and by state assessments.

29. The statement is based on the fact that the correlations are similar. No statistical tests were performed on the differences.

Table 8. Mathematics achievement gains in percentage meeting the primary standard in grade 4, by state: 2000 and 2003

State/ jurisdiction	State			NAEP		
	2000	2003	Gain	2000	2003	Gain
Alabama	57.4	53.1	-4.3	41.7	36.7	-5.0
Alaska	—	67.2	—	—	67.2	—
Arizona	40.6	49.8	9.2	40.6	54.5	13.9 *
Arkansas	35.6	60.1	24.5	35.7	52.5	16.8 *
California	30.9	—	—	30.9	—	—
Colorado	—	86.0	—	—	86.0	—
Connecticut	63.7	58.6	-5.1	63.7	71.7	8.0 *
Delaware	—	72.7	—	—	72.6	—
District of Columbia	36.9	—	—	36.9	—	—
Florida	—	55.4	—	—	55.4	—
Georgia	61.6	73.9	12.3	61.7	76.7	15.0
Hawaii	64.8	67.4	2.6	64.8	78.7	13.9 *
Idaho	—	76.6	—	—	76.7	—
Illinois	52.7	68.3	15.6	52.8	64.7	11.9
Indiana	—	—	—	—	—	—
Iowa	—	75.2	—	—	75.2	—
Kansas	59.3	73.7	14.4	59.3	70.8	11.5
Kentucky	31.1	37.9	6.8	31.1	40.8	9.7
Louisiana	11.3	15.5	4.2	11.4	19.4	8.0 *
Maine	23.3	29.1	5.8	23.3	34.8	11.5 *
Maryland	45.6	—	—	45.6	—	—
Massachusetts	41.0	38.0	-3.0	40.9	50.6	9.7 *
Michigan	76.9	—	—	77.1	—	—
Minnesota	47.9	58.2	10.3	47.9	57.2	9.3
Mississippi	—	74.0	—	—	73.9	—
Missouri	36.6	36.7	0.1	36.6	45.8	9.2 *
Montana	—	75.3	—	—	75.4	—
Nebraska	60.0	—	—	60.0	—	—
Nevada	—	—	—	—	—	—
New Hampshire	—	80.3	—	—	80.3	—
New Jersey	—	67.6	—	—	67.5	—
New Mexico	—	42.0	—	—	42.0	—
New York	67.8	78.8	11.0	67.8	76.4	8.6
North Carolina	84.6	92.2	7.6	84.5	93.2	8.7
North Dakota	—	59.0	—	—	59.0	—
Ohio	42.3	59.0	16.7	42.5	52.9	10.4 *
Oklahoma	85.8	69.3	-16.5	85.9	90.3	4.4 *
Oregon	67.2	77.8	10.6	67.2	81.0	13.8
Pennsylvania	—	56.8	—	—	56.8	—
Rhode Island	20.7	41.8	21.1	20.7	25.6	4.9 *
South Carolina	22.9	32.6	9.7	23.0	38.5	15.5 *
South Dakota	—	72.5	—	—	72.5	—
Tennessee	—	54.5	—	—	41.9	—
Texas	88.5	—	—	88.5	—	—
Utah	50.2	47.9	-2.3	52.9	51.3	-1.6
Vermont	47.0	52.9	5.9	46.9	61.2	14.3 *
Virginia	—	—	—	—	—	—
Washington	—	54.0	—	—	54.0	—
West Virginia	—	—	—	—	—	—
Wisconsin	71.9	—	—	72.1	—	—
Wyoming	25.6	36.4	10.8	25.6	41.6	16.0 *
Average gain	†	†	7.0	†	†	9.9

— Not available.
† Not applicable.
* State and NAEP gains are significantly different from each other at $p<.05$.
NOTE: Primary standard is the state's standard for *proficient* performance.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 and 2003 Mathematics Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

Table 9. Mathematics achievement gains in percentage meeting the primary standard in grade 8, by state: 2000 and 2003

State/ jurisdiction	State			NAEP		
	2000	2003	Gain	2000	2003	Gain
Alabama	55.4	50.7	-4.7	41.1	39.0	-2.1
Alaska	—	65.2	—	—	65.2	—
Arizona	18.2	20.6	2.4	18.2	18.0	-0.2
Arkansas	13.3	21.9	8.6	13.3	18.1	4.8 *
California	23.0	—	—	23.0	—	—
Colorado	—	68.2	—	—	68.2	—
Connecticut	57.2	56.0	-1.2	57.2	60.1	2.9 *
Delaware	—	48.2	—	—	48.3	—
District of Columbia	8.6	—	—	8.5	—	—
Florida	—	54.1	—	—	54.2	—
Georgia	54.8	66.4	11.6	54.9	60.3	5.4 *
Hawaii	60.5	54.1	-6.4	60.5	65.4	4.9 *
Idaho	—	52.5	—	—	52.5	—
Illinois	45.8	53.6	7.8	45.8	48.7	2.9 *
Indiana	—	—	—	—	—	—
Iowa	—	71.7	—	—	71.7	—
Kansas	55.3	59.3	4.0	55.3	55.6	0.3
Kentucky	26.0	31.9	5.9	26.0	29.9	3.9
Louisiana	7.3	8.7	1.4	7.3.0	10.5	3.2
Maine	20.6	17.1	-3.5	20.6	18.8	-1.8
Maryland	51.1	—	—	51.1	—	—
Massachusetts	33.7	38.1	4.4	33.7	41.1	7.4 *
Michigan	—	—	—	—	—	—
Minnesota	—	—	—	—	—	—
Mississippi	—	46.0	—	—	46.1	—
Missouri	12.9	13.4	0.5	12.9	17.9	5.0 *
Montana	—	70.3	—	—	70.4	—
Nebraska	59.7	—	—	60.9	—	—
Nevada	—	—	—	—	—	—
New Hampshire	—	—	—	—	—	—
New Jersey	—	56.2	—	—	56.1	—
New Mexico	—	39.4	—	—	39.3	—
New York	41.4	54.0	12.6	41.4	47.8	6.4 *
North Carolina	80.7	82.2	1.5	80.8	84.1	3.3
North Dakota	—	43.6	—	—	43.6	—
Ohio	—	—	—	—	—	—
Oklahoma	70.5	71.3	0.8	70.5	73.5	3.0
Oregon	54.7	57.6	2.9	54.5	54.6	0.1
Pennsylvania	—	51.5	—	—	51.6	—
Rhode Island	26.5	35.3	8.8	26.6	28.7	2.1 *
South Carolina	19.3	20.2	0.9	19.3	28.4	9.1 *
South Dakota	—	57.5	—	—	57.5	—
Tennessee	—	57.7	—	—	42.6	—
Texas	89.8	—	—	89.9	—	—
Utah	57.9	56.8	-1.1	52.6	53.9	1.3
Vermont	45.6	51.7	6.1	45.7	52.4	6.7
Virginia	—	—	—	—	—	—
Washington	—	36.4	—	—	36.4	—
West Virginia	—	—	—	—	—	—
Wisconsin	42.9	—	—	43.0	—	—
Wyoming	31.5	35.1	3.6	31.6	41.3	9.7 *
Average gain	†	†	3.0	†	†	3.6

— Not available.
† Not applicable.
* State and NAEP gains are significantly different from each other at $p<.05$.
NOTE: Primary standard is the state's standard for *proficient* performance.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 and 2003 Mathematics Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

SUMMARY

Comparisons are made between NAEP and state assessment mathematics achievement trends from 2000 to 2003. Achievement trends are measured by both NAEP and state assessments as gains in school-level percentages meeting the state's primary standard. Comparisons are based on the NAEP sample schools for which we also have state assessment scores. Trend data are available for 37 states. However, in eight of the states for which scores are available, the assessment and/or performance standards were changed during the period between 2000 and 2003; therefore, these states are not included in the trend analysis. As a result, comparisons of mathematics achievement trends from 2000 to 2003 are possible in 24 states for grade 4 and 22 states for grade 8.

In aggregate, in grade 4 but not grade 8, mathematics achievement gains from 2000 to 2003 reported by NAEP are significantly larger than those measured by state assessments. Gains measured by state assessments vary substantially between states; however, variability of gains measured by NAEP between states is only about two-thirds as large as the variation in state assessment results.

NAEP and state assessment gains in grade 4 mathematics are correlated positively. This indicates that, in grade 4 mathematics, states in which state assessments found the largest gains in percentages of students meeting the primary state standard tend to be the states in which NAEP results indicated the largest gains. However, grade 4 mathematics is the only grade and subject that showed a noticeable relationship between the NAEP gains and state assessment gains.

When comparisons between NAEP and state assessment 2000-2003 mathematics achievement trends are made for each state, significant differences are found in 14 out of the 24 states in grade 4 and 11 out of the 22 states in grade 8.