

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 report, our comparison of changes in reading achievement is limited to changes in achievement between the 1997-1998 and 2002-2003 school years (i.e., between 1998 and 2003), between the 1997-1998 and 2001-2002 school years (i.e., between 1998 and 2002), and between the 2001-2002 and 2002-2003 school years (i.e., between 2002 and 2003). For research purposes, analysts may wish to examine trends in earlier NAEP years (e.g., 1993-1994), 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 change analysis.²⁶ 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 excluded scores in particular years from the analysis because they changed their state assessments and/or primary standards during the trend periods; changes in percentages meeting the primary standards in that period will not reflect

26. 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.

their actual changes in achievement. These states are listed below, with the reasons for their exclusion:

- California: changed assessment in 2003; 2003 scores are not used for comparisons with prior years' scores.
- Colorado: rescaled assessment in 2002. Since the state did not participate in State NAEP in 2002, this state is excluded from the trend analysis.
- Indiana: changed assessment in 2003; 2003 scores are not used for comparisons with prior years' scores. Because 1998 scores are not available, this state is excluded from the trend analysis.
- Maine: set the 1998-1999 school year as the base year for achievement trends. 1998 scores are not used for comparisons with 2002 and 2003 scores.
- Maryland: changed assessment in 2003; 2003 scores are not used for comparisons with prior years' scores.
- Michigan: changed assessment in 2003; 2003 scores are not used for comparisons with prior years' scores. Because 1998 scores are not available, this state is excluded from the trend analysis.
- Nevada: changed assessment in 2003; 2003 scores are not used for comparisons with prior years' scores. Because 2002 scores are not available, this state is excluded from the trend analysis.
- Texas: changed assessment in 2003; 2003 scores are not used for comparisons with prior years' scores.
- Virginia: changes performance standards every year and no longitudinal equating from year to year done; this state is excluded from the trend analysis.
- Wisconsin: set new performance standards in 2003; 2003 scores are not used for comparisons with prior years' scores.

It is important to note that achievement changes in percentages 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, achievement trends over time on NAEP and the state assessments across states in terms of percentages achieving the state primary standards.²⁷ 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 year (i.e., 1998 or 2002). In that year, the NAEP and the state assessment percentages match by definition.²⁸

Table 7. Reading achievement gains in percentage meeting the primary state standard in grades 4 and 8: 1998, 2002, and 2003

Statistic	1998 to 2003		1998 to 2002		2002 to 2003	
	State	NAEP	State	NAEP	State	NAEP
Grade 4						
Sample size	8		11		31	
Average gain	8.9	1.6 *	7.8	3.5 *	1.1	-0.4
standard error	1.66	2.29	1.82	2.42	1.60	2.13
Between-state standard deviation	8.1	2.9	4.6	2.7	3.6	2.1
standard error	0.65	0.83	0.56	0.79	0.30	0.38
Grade 8						
Sample size	6		10		29	
Average gain	4.9	-1.6 *	5.9	-1.5 *	0.4	-1.2
standard error	1.18	2.10	1.46	2.40	1.31	2.05
Between-state standard deviation	2.9	2.2	5.5	3.6	3.8	2.5
standard error	0.52	1.06	0.48	0.86	0.32	0.38

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

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), 1998, 2000 and 2003 Reading Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

Averaged over the states in which gains from 1998 to 2003 could be compared, reading achievement gains reported by the state assessment are larger than those in NAEP in both grades. From 1998 to 2002, the gains reported by state assessments for both grades are also significantly larger than those in NAEP. On the other hand, the average of the differences in one-year gains from 2002 to 2003 is not statistically significant (based on approximately 30 states).

There are many possible explanations for the differences in gains measured by NAEP and state assessments from 1998 to 2002 and 2003. In addition to the possibility that states have implemented and been teaching instructional curricula that are more aligned with the content of the particular state's reading assessments than with the NAEP content, there are a number of explanations associated with differences in testing.

27. For Alabama, Tennessee, and Utah, we did not have percentages meeting the state standard; instead, we report changes in percentile ranks.

28. The state profile section of this report compares trends for multiple state standards, not only the single primary standard.



We have constructed the comparisons between NAEP and state assessment results so as to remove two important sources of error, by comparing trends for the same sets of schools and at the same standard level.²⁹ 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; (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 two to eight percentage points. However, these differences may overestimate the actual variation in gains in different states. The standard deviation of gains between states is only about half as large for grade 4 and two-thirds as large for grade 8 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.

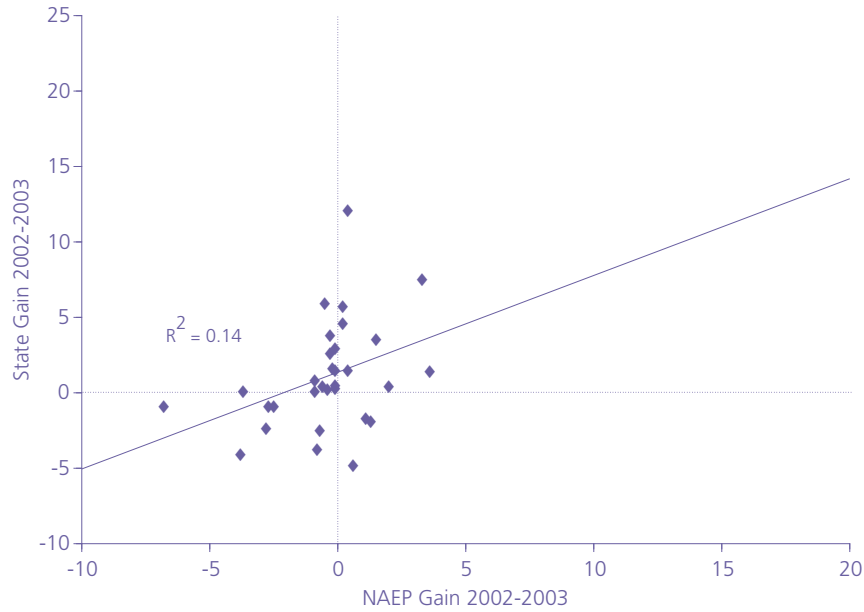
It is possible that the greater variation between states in state assessment gains than in NAEP gains is due to different states' measurement of unique aspects of reading achievement not fully addressed by NAEP. However, an alternative hypothesis must be considered before searching for the unique aspects of reading 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.

To further investigate whether the NAEP and state assessment gains are related, Figures 17 and 18 present scatter plots between the NAEP and state assessment gains from 2002 to 2003 for grades 4 and 8.³⁰ These results 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.

29. Data from the same schools are used to compute both NAEP and state assessment trends. However, different samples of schools are involved in the comparison in different years.

30. For the 1998 to 2003 gains the R^2 was of .05 (8 states) for grade 4 and .06 (6 states) for grade 8. For the 1998 to 2002 gains the R^2 was of .01 (12 states) for grade 4 and .02 (10 states) for grade 8.

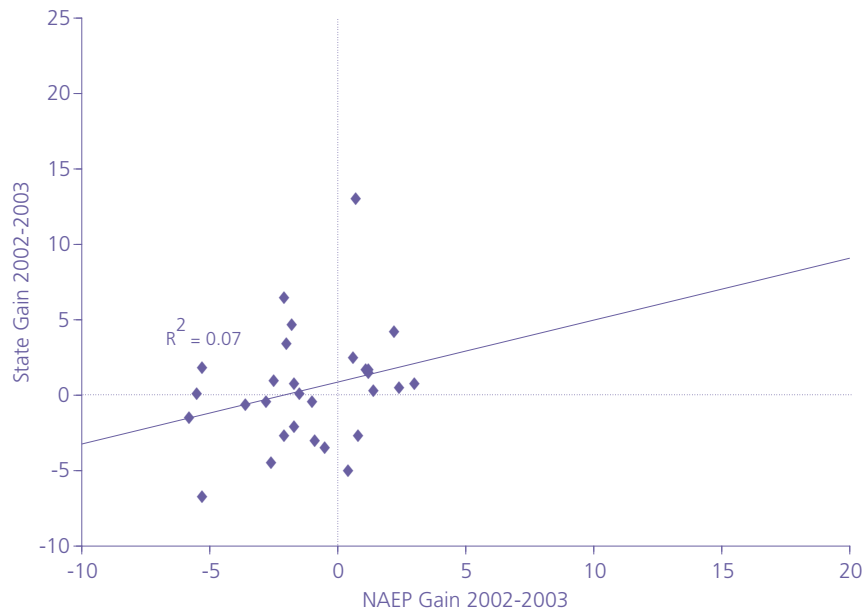
Figure 17. Relationship between NAEP and state assessment gains in percentage meeting the primary state grade 4 reading standard: 2002 to 2003



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), 2002 and 2003 Reading 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 reading standard: 2002 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), 2002 and 2003 Reading Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

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 reading achievement trends are shown in tables 8 and 9.

Table 8 summarizes state-by-state trends for grade 4 reading achievement. From 1998 to 2003, in five out of eight comparisons, there were statistically significantly larger increases in percentages meeting the primary standard on the state assessments than on NAEP; these states are Minnesota, North Carolina, Oregon, Rhode Island, and Washington. Between 1998 and 2002, 5 out of 11 states showed a significantly larger increase on the state assessment than on NAEP; these are Minnesota, Oregon, Rhode Island, Washington, and Wisconsin. When we examine the one-year gains, 6 out of 31 states showed significantly larger reading achievement gains in the state assessment results than in NAEP from 2002 to 2003; these are Arkansas, Florida, Kansas, Kentucky, Minnesota, and North Carolina. It should also be noted that in Massachusetts there was also a statistically significant difference: the percentage meeting the state's primary standard as measured by NAEP decreased by seven percent, compared to a one-percent decrease on the state assessment. On the other hand, between 2002 and 2003, the Connecticut state assessment found about a four percent decrease and the Louisiana state assessment found a five percent decrease in percentage mastering reading skills, while NAEP found significantly smaller changes (of less than one percent) in both states.

Table 9 summarizes state-by-state trends for grade 8 reading scores. Between 1998 and 2003, state assessments found larger reading achievement gains between 1998 and 2003 than NAEP in five out of six states: Connecticut, North Carolina, Oregon, Rhode Island, and Washington. Four of these five states are the same states in which grade 4 results from NAEP and state assessments differed. Between 1998 and 2002, in 6 out of 10 states, state assessments found greater improvements in the percentages meeting the primary standards than NAEP did; these states are California, North Carolina, Oregon, Rhode Island, Texas and Wisconsin. In Connecticut there was also a statistically significant difference in 1998-2002 trends: the percentage meeting the state's primary standard as measured by NAEP decreased by four percent, compared to an increase of over one percent on the state assessment. Between 2002 and 2003, state assessments in six out of 29 states measured statistically larger reading achievement gains than NAEP did; these are Arkansas, Kansas, Maine, Mississippi, Pennsylvania, and Washington. It should be noted that in Delaware and West Virginia, there was also a statistically significant difference: the percentage meeting the state's primary standard as measured by NAEP decreased by about six percent from 2002 to 2003 in both of the states, compared to a decrease of one and a half percent on the state assessment in Delaware and no change in West Virginia. On the other hand, in Louisiana and South Carolina the state assessments measured significantly larger losses from 2002 to 2003 in the percentage meeting the state primary standard than NAEP did.

Table 8. Reading achievement gains in percentage meeting the primary standard in grade 4, by state: 1998, 2002, and 2003

State/ jurisdiction	State						NAEP					
	1998	2002	2003	Gain 98-03	Gain 98-02	Gain 02-03	1998	2002	2003	Gain 98-03	Gain 98-02	Gain 02-03
Alabama	—	53.4	53.7	—	—	0.3	—	42.6	42.5	—	—	-0.1
Alaska	—	—	73.2	—	—	—	—	—	73.1	—	—	—
Arizona	—	56.2	57.6	—	—	1.4	—	56.2	59.8	—	—	3.6
Arkansas	—	56.3	62.0	—	—	5.7	—	56.3	56.5	—	—	0.2 *
California	42.3	48.7	—	—	6.4	—	42.4	48.6	—	—	6.2	—
Colorado	—	—	86.1	—	—	—	—	—	84.8	—	—	—
Connecticut	57.3	58.8	55.0	-2.3	1.5	-3.8	58.0	58.7	57.9	-0.1	0.7	-0.8 *
Delaware	—	78.4	79.9	—	—	1.5	—	78.3	78.7	—	—	0.4
District of Columbia	33.1	—	—	—	—	—	33.1	—	—	—	—	—
Florida	—	53.0	60.5	—	—	7.5	—	53.0	56.3	—	—	3.3 *
Georgia	—	79.4	79.9	—	—	0.5	—	79.3	79.2	—	—	-0.1
Hawaii	—	52.0	50.1	—	—	-1.9	—	44.4	45.7	—	—	1.3
Idaho	—	—	74.4	—	—	—	—	—	74.4	—	—	—
Illinois	—	59.3	59.7	—	—	0.4	—	59.3	58.7	—	—	-0.6
Indiana	—	—	—	—	—	—	—	—	—	—	—	—
Iowa	—	—	75.2	—	—	—	—	—	75.2	—	—	—
Kansas	—	62.4	68.3	—	—	5.9	—	62.5	62.0	—	—	-0.5 *
Kentucky	—	57.7	61.5	—	—	3.8	—	57.6	57.3	—	—	-0.3 *
Louisiana	—	19.3	14.5	—	—	-4.8	—	19.3	19.9	—	—	0.6 *
Maine	—	49.4	49.6	—	—	0.2	—	49.5	49.1	—	—	-0.4
Maryland	37.5	40.8	—	—	3.3	—	37.6	39.9	—	—	2.3	—
Massachusetts	—	54.8	53.9	—	—	-0.9	—	54.8	48.0	—	—	-6.8 *
Michigan	—	56.6	—	—	—	—	—	56.5	—	—	—	—
Minnesota	36.7	49.0	61.1	24.4	12.3	12.1	36.6	37.8	38.2	1.6 *	1.2 *	0.4 *
Mississippi	—	83.3	86.8	—	—	3.5	—	83.4	84.9	—	—	1.5
Missouri	—	34.2	32.5	—	—	-1.7	—	34.2	35.3	—	—	1.1
Montana	—	77.9	77.0	—	—	-0.9	—	78.1	75.6	—	—	-2.5
Nebraska	—	—	78.7	—	—	—	—	—	78.7	—	—	—
Nevada	—	—	48.3	—	—	—	—	—	16.9	—	—	—
New Hampshire	72.1	—	76.6	4.5	—	—	72.1	—	72.4	0.3	—	—
New Jersey	—	—	77.2	—	—	—	—	—	77.2	—	—	—
New Mexico	—	—	44.4	—	—	—	—	—	44.3	—	—	—
New York	—	62.1	63.7	—	—	1.6	—	62.0	61.8	—	—	-0.2
North Carolina	70.2	76.2	80.8	10.6	6.0	4.6	70.3	77.4	77.6	7.3 *	7.1	0.2 *
North Dakota	—	—	75.2	—	—	—	—	—	75.3	—	—	—
Ohio	—	67.6	68.4	—	—	0.8	—	67.7	66.8	—	—	-0.9
Oklahoma	—	—	72.5	—	—	—	—	—	72.5	—	—	—
Oregon	64.5	77.5	77.6	13.1	13.0	0.1	64.5	71.1	67.4	2.9 *	6.6 *	-3.7
Pennsylvania	—	55.9	58.5	—	—	2.6	—	55.9	55.6	—	—	-0.3
Rhode Island	51.2	61.9	61.0	9.8	10.7	-0.9	51.4	51.2	48.5	-2.9 *	-0.2 *	-2.7
South Carolina	—	33.0	30.5	—	—	-2.5	—	33.0	32.3	—	—	-0.7
South Dakota	—	—	85.6	—	—	—	—	—	85.7	—	—	—
Tennessee	—	56.6	54.2	—	—	-2.4	—	47.7	44.9	—	—	-2.8
Texas	85.8	91.9	—	—	6.1	—	85.8	90.9	—	—	5.1	—
Utah	45.6	47.2	47.3	1.7	1.6	0.1	52.2	54.7	53.8	1.6	2.5	-0.9
Vermont	—	73.5	75.0	—	—	1.5	—	73.7	73.6	—	—	-0.1
Virginia	—	—	—	—	—	—	—	—	—	—	—	—
Washington	55.5	68.7	64.6	9.1	13.2	-4.1	55.6	61.2	57.4	1.8 *	5.6 *	-3.8
West Virginia	—	61.1	64.0	—	—	2.9	—	61.1	61.0	—	—	-0.1
Wisconsin	70.8	82.7	—	—	11.9	—	70.9	71.8	—	—	0.9 *	—
Wyoming	—	43.4	43.8	—	—	0.4	—	43.5	45.5	—	—	2.0
Average gain	†	†	†	8.9	7.8	1.1	†	†	†	1.6 *	3.5 *	-0.4

— Not available.

† Not applicable.

* NAEP gains are significantly different from gains reported by the state assessment 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), 1998, 2002, and 2003 Reading Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

Table 9. Reading achievement gains in percentage meeting the primary standard in grade 8, by state: 1998, 2002, and 2003

State/ jurisdiction	State						NAEP					
	1998	2002	2003	Gain 98-03	Gain 98-02	Gain 02-03	1998	2002	2003	Gain 98-03	Gain 98-02	Gain 02-03
Alabama	—	50.7	51.0	—	—	0.3	—	42.2	43.6	—	—	1.4
Alaska	—	—	70.5	—	—	—	—	—	70.5	—	—	—
Arizona	—	55.8	53.7	—	—	-2.1	—	55.9	54.2	—	—	-1.7
Arkansas	—	30.4	43.4	—	—	13.0	—	30.3	31.0	—	—	0.7 *
California	45.2	48.0	—	—	2.8	—	45.2	42.5	—	—	-2.7 *	—
Colorado	—	—	87.3	—	—	—	—	—	87.3	—	—	—
Connecticut	65.0	65.6	68.1	3.1	0.6	2.5	69.7	65.5	66.1	-3.6 *	-4.2 *	0.6
Delaware	—	71.3	69.8	—	—	-1.5	—	71.5	65.7	—	—	-5.8 *
District of Columbia	20.5	—	—	—	—	—	21.1	—	—	—	—	—
Florida	—	47.4	46.8	—	—	-0.6	—	47.4	43.8	—	—	-3.6
Georgia	—	81.3	80.9	—	—	-0.4	—	81.3	80.3	—	—	-1.0
Hawaii	—	54.2	51.5	—	—	-2.7	—	56.5	54.4	—	—	-2.1
Idaho	—	—	73.2	—	—	—	—	—	73.2	—	—	—
Illinois	—	68.6	64.1	—	—	-4.5	—	68.4	65.8	—	—	-2.6
Indiana	—	—	—	—	—	—	—	—	—	—	—	—
Iowa	—	—	68.8	—	—	—	—	—	68.8	—	—	—
Kansas	—	65.1	68.5	—	—	3.4	—	65.2	63.2	—	—	-2.0 *
Kentucky	—	56.4	57.9	—	—	1.5	—	56.5	57.7	—	—	1.2
Louisiana	—	17.8	15.1	—	—	-2.7	—	17.8	18.6	—	—	0.8 *
Maine	—	43.7	44.7	—	—	1.0	—	43.8	41.3	—	—	-2.5 *
Maryland	25.1	22.6	—	—	-2.5	—	25.1	18.2	—	—	-6.9	—
Massachusetts	—	63.8	65.5	—	—	1.7	—	63.7	64.8	—	—	1.1
Michigan	—	53.0	—	—	—	—	—	52.9	—	—	—	—
Minnesota	—	—	—	—	—	—	—	—	—	—	—	—
Mississippi	—	49.1	55.6	—	—	6.5	—	49.0	46.9	—	—	-2.1 *
Missouri	—	31.5	33.2	—	—	1.7	—	31.5	32.7	—	—	1.2
Montana	—	72.3	71.9	—	—	-0.4	—	72.3	69.5	—	—	-2.8
Nebraska	—	—	75.0	—	—	—	—	—	74.9	—	—	—
Nevada	—	—	—	—	—	—	—	—	—	—	—	—
New Hampshire	—	—	—	—	—	—	—	—	—	—	—	—
New Jersey	—	—	73.3	—	—	—	—	—	73.4	—	—	—
New Mexico	—	—	43.9	—	—	—	—	—	43.9	—	—	—
New York	—	42.7	46.9	—	—	4.2	—	42.8	45.0	—	—	2.2
North Carolina	79.0	85.4	85.5	6.5	6.4	0.1	79.3	80.3	78.8	-0.5 *	1.0 *	-1.5
North Dakota	—	—	68.9	—	—	—	—	—	68.9	—	—	—
Ohio	—	—	—	—	—	—	—	—	—	—	—	—
Oklahoma	—	—	78.2	—	—	—	—	—	78.3	—	—	—
Oregon	53.5	65.6	58.9	5.4	12.1	-6.7	53.4	54.0	48.7	-4.7 *	0.6 *	-5.3
Pennsylvania	—	58.5	63.2	—	—	4.7	—	58.6	56.8	—	—	-1.8 *
Rhode Island	36.9	45.2	42.2	5.3	8.3	-3.0	36.9	36.0	35.1	-1.8 *	-0.9 *	-0.9
South Carolina	—	26.1	21.1	—	—	-5.0	—	26.0	26.4	—	—	0.4 *
South Dakota	—	—	78.8	—	—	—	—	—	78.8	—	—	—
Tennessee	—	56.0	56.8	—	—	0.8	—	48.5	46.8	—	—	-1.7
Texas	82.1	94.6	—	—	12.5	—	82.1	79.2	—	—	-2.9 *	—
Utah	51.3	51.2	51.7	0.4	-0.1	0.5	54.3	51.7	54.1	-0.2	-2.6	2.4
Vermont	—	51.9	48.4	—	—	-3.5	—	52.0	51.5	—	—	-0.5
Virginia	—	—	—	—	—	—	—	—	—	—	—	—
Washington	38.7	45.6	47.4	8.7	6.9	1.8	38.7	45.1	39.8	1.1 *	6.4	-5.3 *
West Virginia	—	55.1	55.2	—	—	0.1	—	55.3	49.8	—	—	-5.5 *
Wisconsin	63.6	75.3	—	—	11.7	—	63.5	60.9	—	—	-2.6 *	—
Wyoming	—	38.2	39.0	—	—	0.8	—	38.3	41.3	—	—	3.0
Average gain	†	†	†	4.9	5.9	0.4	†	†	†	-1.6 *	-1.5 *	-1.2

— Not available.

† Not applicable.

* NAEP gains are significantly different from gains reported by the state assessment 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), 1998, 2002, and 2003 Reading Assessments: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.



There are a variety of possible causes for discrepancies between NAEP and state assessment trend results. An obvious explanation would be that some aspects of the state assessment changed during the interval of the trend. Although we omitted trend comparisons for states in which we were aware of such changes, there may have been other changes in procedures of which we were unaware. Another obvious explanation would be that NAEP and the state assessment were testing reading achievement sufficiently differently that trends should not be expected to match. Such differences as test content, testing accommodations, student motivation, and time of testing might be expected to result in lower correlations between NAEP and state assessment results. To test this conjecture, we compared the correlations in 2003 between NAEP and state assessments (the correlations displayed in table 4) in states with significant trend discrepancies (the trends shown in tables 8 and 9) to the correlations in other states.

Overall, there is no noticeable difference between the states with significant discrepancies and those without such discrepancies.³¹ In the grade 4 results, the average correlation between NAEP and state assessments in 2003 is .76 for the states with significant discrepancies in gains from 2002 to 2003 and .70 for the states without discrepancies. Similarly, we did not observe any differences when we examined the states with and without discrepancies in gains from 1998 to 2002 and from 1998 to 2003.

Patterns are very similar in grade 8. The average correlation between NAEP and state assessments is .67 for the states with significant discrepancies in gains from 2002 to 2003 and .73 for the states without discrepancies. When we examined the states with and without discrepancies in gains from 1998 to 2002 and from 1998 to 2003, we did not observe any noticeable differences. These results indicate that in both grades 4 and 8, there is no systematic relationship between: a) the tendency for the two assessments to identify the same schools as low achieving and high achieving, and b) the sizes of discrepancies in gains as measured by NAEP and by state assessments.

SUMMARY

Comparisons are made between NAEP and state assessment reading achievement trends over three trend periods: 1) from 1998 to 2003, 2) from 1998 to 2002, and 3) from 2002 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 36 states. However, in ten of the states for which state assessment scores are available, the assessment and/or the performance standards were changed during the period between 1998 and 2003; therefore, some of these states are not included in the trend analysis for some years.

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

As a result, comparisons of reading achievement trends from 1998 to 2003 are possible in 8 states for grade 4 and 6 states for grade 8, comparisons of trends from 1998 to 2002 are possible in 11 states for grade 4 and 10 states for grade 8, and comparisons of trends from 2002 to 2003 are possible in 31 states for grade 4 and 29 states for grade 8.

When comparisons between NAEP and state assessment 1998-to-2003 reading achievement trends are made for each state, significant differences (higher or lower) are found in five out of the eight states in grade 4 and five out of the six states in grade 8. When we examine trends from 1998 to 2002, significant differences are found in 5 out of the 11 states in grade 4 and seven out of the 10 states in grade 8. As for the trends from 2002 to 2003, significant differences are found in 9 out of the 31 states in grade 4 and 10 out of the 29 states in grade 8.

In aggregate, in both grades 4 and 8, reading achievement gains from 1998 to 2003 reported by state assessments are significantly larger than those measured by NAEP. Between 2002 and 2003, however, the gains are not significantly different between NAEP and state assessments at either grade. For all three time periods, gains measured by state assessments vary substantially between states; however, variability of gains measured by NAEP between states is only about half to two-thirds as large as the variation in state assessment results. At the state level, gains measured by NAEP and state assessments are not significantly correlated with each other. This indicates that the states in which state assessments found the largest gains in percentages of students meeting the primary state reading standard are not necessarily the states in which NAEP results indicated the largest gains.