

# 6

## Supporting Statistics

In this section, we address two of the issues that were faced in preparing to compare NAEP and state assessment results: (1) the changing rates of exclusion and accommodation in NAEP; and (2) the effects of using the NAEP sample of schools for the comparisons.

### STUDENTS WITH DISABILITIES AND ENGLISH LANGUAGE LEARNERS

---

Many factors affect comparisons between NAEP and state assessment measures of reading achievement trends and gaps. One of these factors is the manner in which the assessments treat the problem of measuring the reading achievement of students with disabilities (SD) and English language learners (ELL). Before the 1990s, small percentages of students were excluded from all testing, including national NAEP as well as state assessments. In the 1990s, increasing emphasis was placed on providing equal access to educational opportunities for SD and ELL students, including large-scale testing (Lehr and Thurlow, 2003). Both NAEP and state assessment programs developed policies for accommodating the special testing needs of SD/ELL students to decrease the percentages of students excluded from assessment.

In the period since 1995, NAEP trends haven been subject to variation due to changing exclusion rates in different states (McLaughlin 2000, 2001, 2003). Because that variation confounds comparisons between NAEP and state assessment results, the NAEP computations in this report have been based on full population estimates (FPE). The full population estimates incorporated questionnaire information about the differences between included and excluded SD/ELL students in each state to impute plausible values for the students excluded from the standard NAEP data files (McLaughlin 2000, 2001, 2003; Wise et al., 2004). Selected computations ignoring the subpopulation of students represented by the roughly 5 percent of students excluded from NAEP participation are presented in appendix C. Later in this section, we also compare (in tables 16, 17, 18, and 19) the average differences in the results obtained by using the full population estimates versus those results obtained when we used the standard NAEP data files.

Research on the effects of exclusions and accommodations on assessment results has not yet identified their impact on gaps and trends. However, to facilitate exploration of possible explanations of discrepancies between NAEP and state assessment results in terms of exclusions and accommodations, table 14 displays NAEP estimates of percentages of the population identified, excluded, and accommodated in 1998, 2002, and 2003 for grades 4 and 8.

**Table 14. Percentages of grades 4 and 8 English language learners and students with disabilities identified, excluded, or accommodated in NAEP reading assessments: 1998, 2002, and 2003**

Students	Grade 4			Grade 8		
	1998 <sup>1</sup>	2002 <sup>2</sup>	2003	1998 <sup>1</sup>	2002 <sup>2</sup>	2003
<b>Identified</b>	<b>18.3</b>	<b>20.6</b>	<b>21.9</b>	<b>15.6</b>	<b>17.9</b>	<b>18.5</b>
Students with disabilities	10.5	11.1	11.5	10.3	11.6	12.1
English language learners	7.2	7.8	8.3	4.7	4.8	4.7
Both	0.6	1.7	2.1	0.7	1.5	1.7
<b>Excluded</b>	<b>8.1</b>	<b>7.0</b>	<b>6.3</b>	<b>4.9</b>	<b>5.7</b>	<b>5.2</b>
Students with disabilities	4.4	4.5	3.8	3.4	4.0	3.6
English language learners	3.3	1.6	1.7	1.3	1.1	0.9
Both	0.4	0.8	0.8	0.3	0.6	0.6
<b>Accommodated</b>	<b>2.7</b>	<b>3.7</b>	<b>5.4</b>	<b>2.5</b>	<b>3.7</b>	<b>5.3</b>
Students with disabilities	2.4	3.0	4.3	2.2	3.2	4.5
English language learners	0.3	0.4	0.7	0.2	0.3	0.4
Both	0.1	0.2	0.4	0.1	0.2	0.4

1. Alaska, Idaho, Indiana, North Dakota, Nebraska, New Jersey, Ohio, Pennsylvania, South Dakota, and Vermont are not included in totals for either grade, and Michigan and New Hampshire are not included for grade 8.

2. Alaska, Colorado, New Hampshire, New Jersey, and South Dakota are not included in totals.

NOTE: Detail may not sum to totals because of rounding.

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.

The top segment of table 14 displays the percentages of students identified as SD, ELL, and both in recent NAEP reading assessments. The percentages shown for SD and ELL do not include students identified with both special needs, so each total is the sum of the three subgroups. These percentages include students subsequently excluded from participation, and they are weighted to represent percentages of the student population. The figures in table 14 are aggregated over the states participating in NAEP at the state level in each case.<sup>41</sup> Individual state figures are displayed in the State Profiles section of this report (Appendix D).

The middle segment of table 14 displays the percentages of students who were represented by students who were excluded from participation in the NAEP test sessions. As before, these figures represent percentages of the student population. The

41. For 1998, Alaska, Idaho, Indiana, North Dakota, Nebraska, New Jersey, Ohio, Pennsylvania, South Dakota, and Vermont are not included in totals for either grade, and Michigan and New Hampshire are not included for grade 8. For 2002, Alaska, Colorado, New Hampshire, New Jersey, and South Dakota are not included in totals.

bottom segment of the table displays the percentages of students who were provided with testing accommodations.

Students identified as SD outnumber those identified as ELL by a factor of 3 to 2 in grade 4 and a factor of 5 to 2 at grade 8. There was a 20 percent increase in the aggregate percentage of students identified as either SD or ELL between 1998 and 2003: from 18 to 22 percent at grade 4 and from 16 to 19 percent at grade 8. These percentages and their changes varied substantially between states, as shown in tables in appendix D.

While the figures in table 14 emphasize that the percentages of students who were excluded and accommodated were a small fraction of the students selected to participate in NAEP, they do not show the actual rates of exclusion of those students with disabilities and English language learners. Tables 15 displays these rates, along with the rates at which students with disabilities and English language learners who are included are provided with testing accommodations.

**Table 15. Percentages of those identified as English language learner or as with disabilities, excluded, or accommodated in the NAEP reading assessments grades 4 and 8: 1998, 2002, and 2003**

Students	Grade 4			Grade 8		
	1998 <sup>1</sup>	2002 <sup>2</sup>	2003	1998 <sup>1</sup>	2002 <sup>2</sup>	2003
<b>Excluded</b>	<b>44.5</b>	<b>33.7</b>	<b>28.8</b>	<b>31.6</b>	<b>31.8</b>	<b>28.0</b>
Students with disabilities	42.4	40.3	33.2	32.9	34.8	29.9
English language learners	45.5	20.9	20.1	27.1	22.2	20.1
Both	69.6	50.6	39.0	42.8	40.1	36.3
<b>Accommodated</b>	<b>26.4</b>	<b>26.7</b>	<b>34.7</b>	<b>23.1</b>	<b>30.7</b>	<b>39.7</b>
Students with disabilities	39.1	45.1	55.9	32.2	42.4	52.9
English language learners	6.9	6.8	10.2	5.3	8.7	10.0
Both	28.3	29.3	34.4	17.9	24.8	38.7

1. Alaska, Idaho, Indiana, North Dakota, Nebraska, New Jersey, Ohio, Pennsylvania, South Dakota, and Vermont are not included in totals for either grade, and Michigan and New Hampshire are not included for grade 8.

2. Alaska, Colorado, New Hampshire, New Jersey, and South Dakota are not included in totals.

NOTE: Detail may not sum to totals because of rounding.

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.

At grade 4 in 1998, nearly half of the students identified as SD or ELL were excluded from NAEP, but this percentage was reduced to fewer than 30 percent in 2003 as accommodations were introduced and more widely applied. The reduction was even greater for students identified as ELL. In 1998, a smaller percentage of identified students were excluded in grade 8 than in grade 4; but due to the dramatic decrease at grade 4 between 1998 and 2003, the percentages were similar in 2003 (29 percent in grade 4, 28 percent in grade 8).

NAEP has gradually increased its permission rules and procedures for the use of testing accommodations for SD and ELL in an effort to reduce exclusions. While in 1998 about one quarter of SD and ELL students participating in NAEP sessions were

provided accommodations, by 2003, over a third were provided accommodations. There is little research to address the question of how that increase affects the measurement of trends.

## NAEP FULL POPULATION ESTIMATES AND STANDARD ESTIMATES

---

In this report, unlike previous NAEP reports, achievement estimates based on questionnaire and demographic information for this subpopulation are incorporated in the NAEP results. NAEP statistics presented are based on full population estimates, which include imputed performance for students with disabilities and English language learners who are excluded from participation in NAEP. As shown in table 14, these are roughly 5 percent of the students selected to participate in NAEP. Standard NAEP estimates do not represent this 5 percent of the student population, whose average reading achievement is presumably lower than the reading achievement of the 95 percent of students included in the standard NAEP population estimates. Because the percentages of students excluded by NAEP vary from state to state, from year to year, and between population subgroups, estimates of trends and gaps can be substantially affected by exclusion rates. While we have not been able to adjust for varying exclusion rates in state assessment data in this report, we have, for the most part, eliminated the effects of varying exclusion rates in the NAEP data.

The method of imputation is based on information from a special questionnaire completed for all SDs and ELLs selected for NAEP, whether or not they are excluded. The method of imputation is described in appendix A. The basic assumption of the imputation method is that excluded SDs and ELLs with a particular profile of teacher ratings and demographics would achieve at the same level as the included SDs and ELLs with the same profile of ratings and demographics in the same state.

All comparisons between NAEP and state assessment results in this report were carried out a second time using the standard NAEP estimates. Four tables (tables 16-19) below summarize the comparisons of reading standards, correlations, trends, and gap computations we derived by using the full population estimates (FPE), versus the standard NAEP reported data (SNE). The summary figures in these tables (unweighted averages, standard deviations, and counts of statistically significant differences) are based on the individual state results presented in tables in the preceding sections, which are full population estimates, and on standard NAEP estimates presented in appendix C.

Table 16 below shows the average differences in the NAEP equivalents of primary state reading standards in 2003. Although the FPE-based NAEP equivalents were generally two to three points lower than SNE-based equivalents, due to inclusion of more low-achieving students in the represented population, there was substantial variation between states, due to variations in NAEP exclusion rates between states.

**Table 16. Difference between the NAEP score equivalents of primary reading achievement standards, obtained using full population estimates (FPE) and standard NAEP estimates (SNE), by grade: 2003**

Level	Number of states	Mean difference of NAEP equivalent standards: FPE-SNE	Standard deviation of difference
Grade 4	48	-3.0	1.6
Grade 8	45	-2.3	1.6

NOTE: Primary standard is the state's standard for *proficient* performance. Negative mean differences in the NAEP equivalent standards indicate that the standards based on full population estimates are lower than the standards based on standard NAEP estimates.

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

Table 17 shows the average differences in the correlations (of the NAEP and state assessment percentages meeting grades 4 and 8 reading standards in 2003) between those correlations computed using the full population estimates (presented in table 4) and the correlations computed using the standard NAEP reported data (presented in table C3). The table indicates that there is only a .01 average difference between the correlations based on the full population estimated NAEP data and the correlations obtained using the NAEP reported data.

**Table 17. Difference between correlations of NAEP and state assessment school-level percentages meeting primary state reading standards, obtained using NAEP full population estimates (FPE) and standard NAEP estimates (SNE), by grade: 2003**

Level	Number of states	Mean difference of NAEP equivalent standards: FPE-SNE	Standard deviation of difference
Grade 4	51	0.01	0.03
Grade 8	48	0.01	0.02

NOTE: Primary standard is the state's standard for *proficient* performance. Positive mean differences indicate that the correlations based on the full population estimates are greater than the correlations based on the standard NAEP estimates. For three states, the correlated achievement measure is the median percentile rank.

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

Table 18 shows the average differences in the gains (of 4th and 8th grade reading performance) between those gains computed using the full population estimates (presented in table 7 and table 8) and the trends computed using the standard NAEP estimates (presented in table C4). Although the mean difference in trends is small, the difference varied from state to state and in a few states the difference was sufficient to change the result of a test for statistical significance. In some states NAEP excluded more students in 1998 than in 2003, and in others the reverse was true. Therefore, the effect of exclusions on standard NAEP estimates of trends is to increase the estimate in some states and decrease the estimate in others.

**Table 18. Mean difference in reading performance gains obtained using NAEP full population estimates (FPE) versus standard NAEP estimates (SNE), by grade: 1998, 2002, and 2003**

Level	Interval	Number of states	Mean difference in gain: FPE-SNE		Standard deviation of difference in gain: FPE-SNE		Number of statistically significant differences between NAEP and state assessment gains	
			State	NAEP	State	NAEP	FPE	SNE
Grade 4	1998-2003	8	0.2	0.1	0.4	1.2	5	5
	1998-2002	11	0.5	0.5	1.3	2.0	5	5
	2002-2003	31	0.1	0.0	0.2	0.7	9	8
Grade 8	1998-2003	6	-0.2	0.4	0.1	0.6	5	5
	1998-2002	10	0.2	-0.3	0.7	0.9	7	6
	2002-2003	29	0.1	0.0	0.2	0.6	10	8

NOTE: Positive mean differences in the NAEP equivalent standards indicate that the gains based on full population estimates are larger, though not always significantly, than the gains based on standard NAEP estimates.

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.

One aspect of the results in table 18 deserves comment: the effect of the difference between full population estimates and standard NAEP estimates on the values for the state assessment trends. Because full population estimation increases the weight assigned to schools that exclude students from NAEP in computing state averages, the average state assessment score estimated for a state, based on NAEP schools, changes with the changing weights. For example, if schools that exclude larger percentages of students from NAEP are also schools with lower average state assessment scores, then the state average state assessment score based on weights used in full population estimation will be lower than the average based on the weights used for standard NAEP estimates.

Finally, we compare the differences between full population estimates and standard NAEP estimates on gap comparisons. Table 19 shows the average differences in the achievement gaps (in 4th and 8th grade reading performance) between those gaps computed using the full population estimates (presented in table 10 and table 11) and the achievement gaps computed using the NAEP reported data (presented in table C6). The figures in tables 10 and 11 and C6 are differences between the gaps as measured by NAEP and the gaps as measured by state assessments. A positive entry in those tables indicated that the NAEP measure of the gap was smaller than the state assessment of the gap. For table 19, we subtract the NAEP-state assessment differences based on standard NAEP estimates from the NAEP-state assessment differences based on full population estimates.

Overall, the gap comparison results for full population estimates are similar to the results for standard NAEP estimates. However, at grade 8, there was a significant difference in the Black-White gap in Georgia, only according to the standard NAEP estimates, and there was a significant difference in the Hispanic-White gap in Illinois, only according to the full population estimates. Similarly, significant

differences in the poverty gap were found in California, according to full population estimates, and in Kentucky and South Dakota, according to the standard NAEP estimates.

**Table 19. Mean difference in gap measures of reading performance obtained using NAEP full population estimates (FPE) versus standard NAEP estimates (SNE), by grade: 2003**

Level	Gap	Number of states	Mean difference in gaps: FPE-SNE	Standard deviation of difference in gaps: FPE-SNE	Number of statistically significant differences between NAEP and state assessment gaps	
					FPE	SNE
Grade 4	Black-White	26	1.3	1.1	4	4
	Hispanic-White	14	-0.9	1.0	2	2
	Poverty	31	-0.4	0.8	1	1
Grade 8	Black-White	20	0.6	0.7	2	3
	Hispanic-White	13	-0.9	0.7	3	2
	Poverty	28	-0.8	1.0	9	10

NOTE: Positive mean differences indicate that NAEP finds smaller gaps than state assessments to a greater extent when using the full population estimates than when using standard NAEP estimates.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment: Full population estimates. The National Longitudinal School-level State Assessment Score Database (NLSASD) 2004.

## USE OF SCHOOL-LEVEL DATA FOR COMPARISONS BETWEEN NAEP AND STATE ASSESSMENT RESULTS

One of the critical issues for NAEP-state assessment comparisons is whether the comparisons are based on the same populations. In order to ensure that differences that might be found between NAEP and state assessment results would not be attributable to different sets of schools, our comparisons were carried out on schools in the NAEP sample, and summary state figures were constructed from the results in those schools, using NAEP weights. One barrier to this approach was the challenge of finding the state assessment scores for the several thousand schools participating in each of the NAEP assessments. In this section, we present information on that matching process. In addition, as a validation of both the NAEP sample and the match between (a) the state assessment data on the databases we used and (b) the data used by the states for their reports, we compare our estimates of the percentages of students meeting state standards with the percentages reported on state websites.

### State assessment results for NAEP schools

Our aim was to match state assessment scores to all of the public schools participating in NAEP. The percent of schools matched for the 2003 NAEP assessments are displayed in table 20. At grade 4, the median match rate across states was 99 percent. That is, of the approximately 100 schools per state per assessment, we found state assessment records for all, or all but one, in most states. The fact that the median weighted match rate was 99.6 percent indicates that the schools we missed tended to

be schools carrying less weight in computing state averages from the NAEP sample. The overall success of the matching process was equally good at grade 8, where the median match rate was 99.2 percent, with a median weighted match rate of 99.8 percent.

For grade 4, the only jurisdictions with a matching rate of less than 90 percent were the District of Columbia (87 percent) and South Dakota (89 percent). In South Dakota, some of the unmatched schools are likely to be small schools for which all state assessment scores are suppressed. Schools having all missing data for assessment results in state assessment files had purposefully been excluded from the NLSLSASD, the database from which we extracted state assessment information for this report. These tended to be small schools, which are more prevalent in rural states such as South Dakota. The weighted match rate for South Dakota was 98 percent; that is, matches were found for NAEP schools representing 98 percent of the student population.

For grade 8, we were able to match more than 90 percent of the schools in all but five jurisdictions: District of Columbia (73 percent), Ohio (75 percent), Hawaii (85 percent), New Hampshire (87 percent), and South Dakota (90 percent). For Ohio, we do not include any grade 8 results in this report; for Hawaii, Nebraska, and South Dakota, the weighted match rates are very high. However, for the District of Columbia, the lower match rate may offer one explanation for any discrepancies that are found between NAEP and state assessment results for grade 8.

Failure to match a NAEP school to the state records is not the only source of omission of NAEP schools from the comparison database. As indicated in table 1, the percentages of schools used for analyses were somewhat lower in certain states. In many states, the percentages of the population represented in the analyses clustered around 90 percent; however, the comparison samples in Arizona, Delaware, New Mexico, and Tennessee included schools that represented less than 85 percent of the NAEP sample at grade 4. At grade 8, less than 85 percent of the student population was represented in the analyses for the District of Columbia, New Mexico, North Dakota, Tennessee, and Washington. Failure to match all NAEP schools is not likely to have a significant impact on the comparison analyses unless the missing schools are systematically different from other schools. In fact, due to suppression of state assessment scores for small reporting samples, in these analyses missing schools are more likely to be small schools. Interpretation of the findings should take this potential bias into account.

This is an even more critical issue with respect to the gap analyses, where small to moderate-sized schools with small percentages of minority students are more likely to have their minority average achievement scores suppressed. And to balance the gap analyses, schools with only one or two NAEP minority participants were excluded from the minority population used to construct the population achievement profile for that minority. The percentages of the minorities represented by the NAEP data that are included in gap analyses in each state are displayed in table 21.



Across the states for which gap profiles are included in this report, the median percentages of Black students and disadvantaged students included in grade 4 analyses is 88 percent, and the median percentage of Hispanic students is 85 percent. In most states, more than two-thirds of the minority students are included, and in all states, more than half are included. The states with fewer than two-thirds of Black students included are Connecticut, Delaware, Kansas, and Wisconsin. Connecticut Hispanic gap analyses are based on fewer than two-thirds of the Hispanic students in Connecticut, and poverty gap analyses in New York, Vermont, and New Hampshire are based on fewer than two-thirds of the disadvantaged students in NAEP files.

At grade 8, due to larger schools, fewer minority data are suppressed in state assessment files. The median percentages included in gap analyses are 94 percent for Blacks, 93 percent for Hispanics, and 91 percent for disadvantaged students; in no states are analyses based on fewer than 70 percent of the minority students in NAEP files.

Table 20. Weighted and unweighted percentages of NAEP schools matched to state assessment records in reading, by grade and state: 2003

State/jurisdiction	Grade 4		Grade 8	
	unweighted	weighted	unweighted	weighted
Alabama	99.1	97.5	99.0	98.9
Alaska	100.0	100.0	100.0	100.0
Arizona	93.3	91.0	96.6	96.0
Arkansas	100.0	100.0	100.0	100.0
California	98.8	98.9	99.5	99.9
Colorado	96.8	97.0	97.4	98.4
Connecticut	99.1	99.9	100.0	100.0
Delaware	92.0	92.5	97.3	98.3
District of Columbia	87.3	88.8	73.0	83.0
Florida	98.1	98.1	99.0	99.4
Georgia	96.2	96.5	96.6	95.3
Hawaii	100.0	100.0	84.8	99.0
Idaho	100.0	100.0	100.0	100.0
Illinois	99.4	99.6	100.0	100.0
Indiana	100.0	100.0	100.0	100.0
Iowa	97.8	98.0	98.3	98.3
Kansas	99.3	97.9	99.2	99.4
Kentucky	100.0	100.0	100.0	100.0
Louisiana	100.0	100.0	100.0	100.0
Maine	98.7	99.8	98.2	99.8
Maryland	100.0	100.0	100.0	100.0
Massachusetts	100.0	100.0	100.0	100.0
Michigan	99.3	99.6	100.0	100.0
Minnesota	99.1	99.8	99.1	100.0
Mississippi	99.1	98.6	100.0	100.0
Missouri	100.0	100.0	100.0	100.0
Montana	99.5	99.9	100.0	100.0
Nebraska	90.4	97.1	91.2	98.7
Nevada	98.2	95.0	97.0	97.0
New Hampshire	99.2	98.8	86.9	85.7
New Jersey	100.0	100.0	100.0	100.0
New Mexico	98.3	98.4	93.8	94.7
New York	98.0	98.0	97.3	98.0
North Carolina	98.7	99.6	98.5	98.6
North Dakota	98.6	99.8	99.3	99.9
Ohio	98.2	93.5	75.2	67.3
Oklahoma	100.0	100.0	99.2	99.5
Oregon	99.2	98.8	100.0	100.0
Pennsylvania	90.4	90.8	100.0	100.0
Rhode Island	99.1	99.7	96.4	98.3
South Carolina	97.2	98.3	99.0	98.2
South Dakota	88.8	98.2	89.8	98.5
Tennessee	100.0	100.0	99.1	96.7
Texas	99.0	97.9	98.6	94.8
Utah	98.3	97.6	100.0	100.0
Vermont	99.4	97.8	100.0	100.0
Virginia	95.7	93.8	96.3	93.0
Washington	99.1	99.9	100.0	100.0
West Virginia	100.0	100.0	100.0	100.0
Wisconsin	100.0	100.0	100.0	100.0
Wyoming	98.8	99.6	95.5	99.6

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

Table 21. Percentages of NAEP student subpopulations in grades 4 and 8 included in comparison analysis for reading, by state: 2003

State/ jurisdiction	Black students		Hispanic students		Disadvantaged students	
	Grade 4	Grade 8	Grade 4	Grade 8	Grade 4	Grade 8
Alabama	92.8	89.5	—	—	92.7	89.1
Alaska	—	—	—	—	—	—
Arizona	—	—	79.3	89.1	—	—
Arkansas	93.6	83.6	—	—	97.6	91.7
California	—	—	94.7	98.5	95.5	98.2
Colorado	—	—	—	—	—	—
Connecticut	55.8	76.3	60.1	75.6	75.1	83.7
Delaware	56.8	95.7	—	—	67.0	96.6
District of Columbia	—	—	—	—	70.9	77.8
Florida	95.2	99.5	88.2	97.4	97.0	97.2
Georgia	89.7	95.0	—	—	92.5	96.0
Hawaii	—	—	—	—	92.3	96.3
Idaho	—	—	70.9	73.3	—	—
Illinois	81.0	93.1	88.7	94.5	81.0	91.1
Indiana	84.6	92.9	—	—	90.8	99.4
Iowa	—	—	—	—	—	—
Kansas	57.2	—	—	—	82.0	86.9
Kentucky	81.6	—	—	—	98.9	97.7
Louisiana	95.5	98.1	—	—	98.9	96.9
Maine	—	—	—	—	—	—
Maryland	—	—	—	—	—	—
Massachusetts	75.8	—	71.9	—	—	—
Michigan	—	—	—	—	—	—
Minnesota	—	—	—	—	87.9	—
Mississippi	93.7	93.0	—	—	91.2	88.8
Missouri	70.5	79.8	—	—	—	—
Montana	—	—	—	—	—	—
Nebraska	—	—	—	—	—	—
Nevada	72.0	96.2	91.6	96.6	79.3	84.6
New Hampshire	—	—	—	—	59.0	—
New Jersey	88.0	95.8	84.9	90.7	86.7	96.0
New Mexico	—	—	70.7	74.2	69.8	86.8
New York	79.9	81.6	84.1	78.9	62.8	70.0
North Carolina	94.8	97.1	—	—	95.1	97.3
North Dakota	—	—	—	—	—	—
Ohio	87.9	—	—	—	84.6	—
Oklahoma	95.6	—	—	—	—	—
Oregon	—	—	—	91.7	—	—
Pennsylvania	76.2	94.0	—	—	79.2	97.2
Rhode Island	—	—	86.4	95.9	—	—
South Carolina	92.0	83.7	—	—	94.0	87.8
South Dakota	—	—	—	—	72.7	77.1
Tennessee	97.9	84.9	—	—	97.1	80.7
Texas	92.2	95.4	97.1	96.5	—	—
Utah	—	—	—	—	—	—
Vermont	—	—	—	—	61.7	74.1
Virginia	88.1	97.9	—	—	—	—
Washington	—	—	71.3	—	—	—
West Virginia	—	—	—	—	98.8	83.4
Wisconsin	57.3	—	—	—	87.6	88.3
Wyoming	—	—	—	—	95.0	92.6

— Not available.

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

## STATE ASSESSMENT RESULTS FOR NAEP SAMPLES AND SUMMARY FIGURES REPORTED BY STATES

---

All of the comparisons in this report were based on NAEP and state assessment data for the same schools, weighted by NAEP sampling weights to represent the public school students in the state. Theoretically, the weighted average of the state assessment scores in NAEP schools is an unbiased estimate of state-level statistics. There are several explanations for discrepancies between official state figures and results based on aggregation of state assessment results in the NAEP schools. Suppression of scores in some schools due to small number of students, failure to match state assessment scores to some NAEP schools, inclusion of different categories of schools and students in state figures, and summarization of scores in state reports to facilitate communication, can distort state-level estimates from NAEP schools. Tables 22 and 23 show the percentages of students meeting the primary standard for NAEP samples and states' published reports of reading achievement, for grades 4 and 8 respectively.

There are several reasons for failure to match some NAEP schools. For example, in states in which the only results available to compare to NAEP grade 4 results are grade 3 statistics, there might be a few NAEP schools that serve only grades 4 to 6, and these would have no grade 3 state assessment scores. Similarly, in sampling, NAEP does not cover special situations such as home schooling, and these may be included in state statistics. Finally, in reporting, to be succinct a state may issue reports with single summaries of scores across grades, while the data we analyzed might be specifically grade 4 scores. In fact, because NAEP samples are drawn with great care, factors such as these are more likely sources of discrepancies in tables 22 and 23 than sampling variation.

Table 22. Percentages of grade 4 students meeting primary standard of reading achievement in NAEP samples and states' published reports, by state: 1998, 2002, and 2003

State/ jurisdiction	NAEP			State reports		
	1998	2002	2003	1998	2002	2003
Alabama	—	53.4	53.7	—	—	—
Alaska	—	—	73.2	—	74.6	71.3
Arizona	—	56.2	57.6	—	59.0	57.0
Arkansas	—	56.3	62.0	—	57.0	61.0
California	42.3	48.7	—	40.0	49.0	—
Colorado	—	—	86.1	55.0	61.0	63.0
Connecticut	57.3	58.8	55.0	—	57.9	55.9
Delaware	—	78.4	79.9	—	78.0	78.0
District of Columbia	33.1	—	—	—	—	—
Florida	—	53.0	60.5	—	55.0	60.0
Georgia	—	79.4	79.9	—	79.0	80.0
Hawaii	—	52.0	50.1	—	—	—
Idaho	—	—	74.4	—	—	75.6
Illinois	—	59.3	59.7	—	59.2	60.4
Indiana	—	—	—	68.0	66.0	72.0
Iowa	—	—	75.2	69.8	69.0	70.4
Kansas	—	62.4	68.3	—	63.0	68.9
Kentucky	—	57.7	61.5	—	60.2	62.3
Louisiana	—	19.3	14.5	—	19.0	14.0
Maine	—	49.4	49.6	—	49.0	49.0
Maryland	37.5	40.8	—	—	42.2	—
Massachusetts	—	54.8	53.9	—	54.6	56.0
Michigan	—	56.6	—	58.6	56.8	75.0
Minnesota	36.7	49.0	61.1	—	48.8	59.4
Mississippi	—	83.3	86.8	—	83.7	87.0
Missouri	—	34.2	32.5	—	35.4	34.1
Montana	—	77.9	77.0	—	75.0	76.0
Nebraska	—	—	78.7	—	—	—
Nevada	—	—	48.3	—	—	—
New Hampshire	72.1	—	76.6	34.0	—	37.0
New Jersey	—	—	77.2	—	79.1	—
New Mexico	—	—	44.4	—	—	—
New York	—	62.1	63.7	—	62.0	64.0
North Carolina	70.2	76.2	80.8	—	—	—
North Dakota	—	—	75.2	—	—	—
Ohio	—	67.6	68.4	—	64.0	68.0
Oklahoma	—	—	72.5	—	—	—
Oregon	64.5	77.5	77.6	66.0	79.0	76.0
Pennsylvania	—	55.9	58.5	—	57.0	58.0
Rhode Island	51.2	61.9	61.0	50.5	62.6	62.8
South Carolina	—	33.0	30.5	—	77.1	81.1
South Dakota	—	—	85.6	—	—	—
Tennessee	—	56.6	54.2	—	—	—
Texas	85.8	91.9	—	86.0	92.0	—
Utah	45.6	47.2	47.3	—	—	—
Vermont	—	73.5	75.0	—	74.5	75.5
Virginia	—	—	—	68.0	78.0	—
Washington	55.5	68.7	64.6	55.6	65.6	66.7
West Virginia	—	61.1	64.0	—	—	—
Wisconsin	70.8	82.7	—	69.0	79.0	—
Wyoming	—	43.4	43.8	—	44.0	44.0

— Not available.

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. State reports are from state education agency website.

**Table 23. Percentages of grade 8 students meeting primary standard of reading achievement for NAEP samples and states' published reports, by state: 1998, 2002, and 2003**

State/ jurisdiction	NAEP			State reports		
	1998	2002	2003	1998	2002	2003
Alabama	—	50.7	51.0	—	—	—
Alaska	—	—	70.5	—	81.6	67.9
Arizona	—	55.8	53.7	—	56.0	55.0
Arkansas	—	30.4	43.4	—	32.0	41.0
California	45.2	48.0	—	46.0	49.0	—
Colorado	—	—	87.3	—	65.0	66.0
Connecticut	65.0	65.6	68.1	—	66.3	68.1
Delaware	—	71.3	69.8	—	72.0	70.0
District of Columbia	20.5	—	—	—	—	—
Florida	—	47.4	46.8	—	45.0	49.0
Georgia	—	81.3	80.9	—	80.0	81.0
Hawaii	—	54.2	51.5	—	—	—
Idaho	—	—	73.2	—	—	74.0
Illinois	—	68.6	64.1	—	68.0	63.7
Indiana	—	—	—	75.0	68.0	63.0
Iowa	—	—	68.8	72.1	69.4	69.3
Kansas	—	65.1	68.5	—	66.8	70.6
Kentucky	—	56.4	57.9	—	55.7	57.3
Louisiana	—	17.8	15.1	—	17.0	15.0
Maine	—	43.7	44.7	—	43.0	45.0
Maryland	25.1	22.6	—	—	23.6	—
Massachusetts	—	63.8	65.5	—	64.0	66.0
Michigan	—	53.0	—	48.8	50.9	61.0
Minnesota	—	—	—	—	—	—
Mississippi	—	49.1	55.6	—	48.4	56.7
Missouri	—	31.5	33.2	—	32.0	32.5
Montana	—	72.3	71.9	—	71.0	70.0
Nebraska	—	—	75.0	—	—	—
Nevada	—	—	—	—	—	—
New Hampshire	—	—	—	—	—	—
New Jersey	—	—	73.3	—	73.2	—
New Mexico	—	—	43.9	—	—	—
New York	—	42.7	46.9	—	44.0	45.0
North Carolina	79.0	85.4	85.5	—	85.1	85.7
North Dakota	—	—	68.9	—	—	—
Ohio	—	—	—	—	—	—
Oklahoma	—	—	78.2	—	—	—
Oregon	53.5	65.6	58.9	55.0	64.0	61.0
Pennsylvania	—	58.5	63.2	—	58.8	63.4
Rhode Island	36.9	45.2	42.2	—	43.9	42.3
South Carolina	—	26.1	21.1	—	26.2	19.9
South Dakota	—	—	78.8	—	—	—
Tennessee	—	56.0	56.8	—	—	—
Texas	82.1	94.6	—	85.0	94.0	—
Utah	51.3	51.2	51.7	—	—	—
Vermont	—	51.9	48.4	—	52.0	49.0
Virginia	—	—	—	65.0	69.0	—
Washington	38.7	45.6	47.4	38.4	44.5	47.9
West Virginia	—	55.1	55.2	—	—	—
Wisconsin	63.6	75.3	—	64.0	74.0	—
Wyoming	—	38.2	39.0	—	38.0	39.0

— Not available.

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. State reports are from state education agency website.



## References

Campbell, J.R., Hombro, C.M., and Mazzeo, J. (2000). *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance*. (NCES 2000-469). U.S. Department of Education, Washington, DC: National Center for Education Statistics.

Forgione Jr., Pascal D. (1999), *Issues surrounding the release of the 1998 NAEP reading report card*. Testimony to the Committee on Education and the Workforce, U.S. House of Representatives, on May 27, 1999. Retrieved May 3, 2007, from [http://www.house.gov/ed\\_workforce/hearings/106th/oi/naep52799/forgione.htm](http://www.house.gov/ed_workforce/hearings/106th/oi/naep52799/forgione.htm)

Holland, P.W. (2002). Two measures of change in the gaps between the cdf's of test score distributions. *Journal of Educational and Behavioral Statistics*, 27: 3-17.

Lehr, C., and Thurlow, M. (2003). *Putting it all together: Including students with disabilities in assessment and accountability systems* (Policy Directions No. 16). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from the World Wide Web: <http://education.umn.edu/NCEO/OnlinePubs/Policy16.htm>

McLaughlin, D.H. (2000). *Protecting State NAEP Trends from Changes in SD/LEP Inclusion Rates* (Report to the National Institute of Statistical Sciences). Palo Alto, CA: American Institutes for Research.

McLaughlin, D.H. (2001). *Exclusions and accommodations affect State NAEP gain statistics: Mathematics, 1996 to 2000* (Appendix to chapter 4 in the NAEP Validity Studies Report on Research Priorities). Palo Alto, CA: American Institutes for Research.

McLaughlin, D.H. (2003). *Full-Population estimates of reading gains between 1998 and 2002* (Report to NCES supporting inclusion of full population estimates in the report of the 2002 NAEP reading assessment). Palo Alto, CA: American Institutes for Research.

McLaughlin, D. (2005, December 8-9). *Considerations in using the longitudinal school-level state assessment score database*. Paper submitted to The National Academies Symposium on the Use of School-Level Data for Evaluating Federal Education Programs.

McLaughlin, D., Bandeira de Mello, V., Blankenship, C., Chaney, K., Esra, P., Hikawa, H., Rojas, D., William, P., and Wolman, M. (2008). *Comparison between NAEP and state mathematics assessment results: 2003* (NCES 2008-475). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

National Assessment Governing Board–NAGB (2001). *Using the National Assessment of Educational Progress To Confirm State Test Results*, A Report of The Ad Hoc Committee on Confirming Test Results, March 2001.

National Center for Education Statistics (2005). *NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics: Findings in Brief* (NCES 2005–463). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office

Ravitch, D. (2004). *A Brief History of Testing and Accountability*, Hoover Digest, 2004. Retrieved from the World Wide Web:  
<http://www.hoover.org/publications/digest/4495866.html>

Shkolnik, J., and Blankenship, C. (2006), *School-level achievement data: A comparison of the use of mean scale scores and proficiency rates*, Palo Alto, CA: American Institutes for Research.

U.S. Department of Education (2002), *No Child Left Behind: A Desktop Reference*, Washington, DC: Office of Elementary and Secondary Education.

Vinovskis, M. A.(1998). *Overseeing the Nation's Report Card: The creation and evolution of the National Assessment Governing Board (NAGB)*, Paper prepared for the National Assessment Governing Board (NAGB), November 1998. Retrieved from the World Wide Web <http://www.nagb.org/pubs/95222.pdf>

Wise, L.L., Le, H., Hoffman, R.G., & Becker, D.E. (2004). *Testing NAEP full population estimates for sensitivity to violation of assumptions*. Technical Report TR-04-50. Alexandria, VA: HumRRO.

Wise, L.L., Le, H., Hoffman, R.G., & Becker, D.E. (2006). *Testing NAEP full population estimates for sensitivity to violation of assumptions: Phase II Draft Technical Report DTR-06-08*. Alexandria, VA: HumRRO.