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 mathematics achievement trends and gaps. One of these factors is the manner in which the assessments treat the problem of measuring the mathematics achievement of students with disabilities (SD) and English language learners (ELL). Before the 1990s, small percentages of students were excluded from 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 have 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 14, 15, 16, and 17) the average differences in the results obtained by using the full population estimates versus those results obtained when we used the standard NAEP estimates.

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 12 displays NAEP estimates of percentages of the population identified, excluded, and accommodated in 2000 and 2003 for grades 4 and 8.

Table 12. Percentages of grades 4 and 8 English language learners and students with disabilities identified, excluded, or accommodated in NAEP mathematics assessments: 2000 and 2003

	Grade 4		Grade 8	
Students	**2000[1]**	**2003**	**2000[1]**	**2003**
Identified	**19.0**	**22.2**	**16.9**	**18.5**
Students with disabilities	10.7	11.7	11.2	12.1
English language learners	7.4	8.5	4.9	4.7
Both	0.9	2.0	0.9	1.6
Excluded	**4.2**	**3.9**	**4.4**	**3.8**
Students with disabilities	2.7	2.4	3.2	2.6
English language learners	1.1	0.9	0.8	0.7
Both	0.4	0.6	0.4	0.5
Accommodated	**6.7**	**8.3**	**4.1**	**6.8**
Students with disabilities	4.4	6.1	3.1	5.6
English language learners	2.1	1.5	0.7	0.7
Both	0.2	0.7	0.2	0.5

1. Alaska, Colorado, Delaware, Florida, New Hampshire, New Jersey, Pennsylvania, South Dakota, and Washington are not included in totals, and Iowa is not included for grade 8.

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), 2000 and 2003 Mathematics Assessments.

The top segment of table 12 displays the percentages of students identified as SD, ELL, and both, in recent NAEP mathematics 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 are aggregated over the states participating in NAEP at the state level in each case.[38] Individual state figures are displayed in the State Profiles section of this report (Appendix D).

The middle segment of table 12 displays the percentages of students who were excluded from participation in the NAEP test sessions. As before, these figures represent percentages of the student population. The bottom segment of the table displays the percentages of students who were provided with testing accommodations.

38. For 2000, Alaska, Colorado, Delaware, Florida, New Hampshire, New Jersey, Pennsylvania, South Dakota, and Washington are not included in totals, and Iowa is not included for grade 8.

Students identified as SD outnumber those identified as ELL by a factor of 3 to 2 in grade 4 and a factor of approximately 2 to 1 at grade 8. There was a 10 percent or more increase in the aggregate percentage of students identified as either SD or ELL between 2000 and 2003: from 19 to 22 percent at grade 4 and from 17 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 12 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 students with disabilities and English language learners. Table 13 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 13. Percentages of those identified as English language learner or as with disabilities, excluded, or accommodated in the NAEP mathematics assessments grades 4 and 8: 2000 and 2003

	Grade 4		Grade 8	
Students identified	**2000[1]**	**2003**	**2000[1]**	**2003**
Excluded	**22.3**	**17.6**	**26.0**	**20.3**
Students with disabilities	25.7	20.6	28.3	21.4
English language learners	14.8	10.6	16.8	14.6
Both	44.1	28.8	48.9	28.2
Accommodated	**45.5**	**45.4**	**32.8**	**46.2**
Students with disabilities	54.9	65.2	39.2	58.9
English language learners	33.8	20.4	18.4	18.4
Both	42.0	49.3	47.6	38.6

1. Alaska, Colorado, Delaware, Florida, New Hampshire, New Jersey, Pennsylvania, South Dakota, and Washington are not included in totals, and Iowa is not included for grade 8.

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), 2000 and 2003 Mathematics Assessments.

At grade 4 in 2000, 22 percent of the students identified as SD or ELL were excluded from NAEP, but this percentage was reduced to fewer than 18 percent in 2003. In 2000 and 2003, a smaller percentage of identified students were excluded from mathematics sessions in grade 4 than 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. By 2000, nearly one half of SD and ELL students participating in grade 4 NAEP sessions were provided accommodations, and this remained constant through 2003. However, in grade 8, the percentage increased from about one third to nearly one half between 2000 and 2003. 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 12, these are roughly 4 percent of the students selected to participate in NAEP. Standard NAEP estimates do not represent this 4 percent of the student population, whose average mathematics achievement is presumably lower than the mathematics achievement of the 96 percent of students included in the standard NAEP 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 14-17) below summarize the comparisons of mathematics standards, correlations, trends, and gap computations we derived by using the full population estimates (FPE), versus the standard NAEP estimates (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 standard NAEP estimates presented in appendix C.

Table 14 below shows the average differences in the NAEP equivalents of primary state mathematics standards in 2003. Although the FPE-based NAEP equivalents were about one point lower than SNE-based equivalents, due to inclusion of more low achieving students in the represented population, there was noticeable variation between states, due to variations in NAEP exclusion rates between states

Table 14. Difference between the NAEP score equivalents of primary mathematics 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	46	-1.1	0.7
Grade 8	43	-1.1	0.8

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 Mathematics Assessment: Full population estimates. The National Longitudinal School-level State Assessment Score Database (NLSLSASD) 2004.

Table 15 shows that there were virtually no differences in the correlations (of the NAEP and state assessment percentages meeting grades 4 and 8 mathematics standards in 2003) between those correlations computed using the full population estimates (presented in table 4) and the standard NAEP estimates (in table C3).

Table 15. Difference between correlations of NAEP and state assessment school-level percentages meeting primary state mathematics 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	49	0.00	0.01
Grade 8	46	0.00	0.01

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 Mathematics Assessment: Full population estimates. The National Longitudinal School-level State Assessment Score Database (NLSLSASD) 2004.

Table 16 shows the average differences in the trends of 4th and 8th grade mathematics performance from 2000 to 2003 when those trends are computed (a) using the full population estimates (presented in table 8 and table 9) and (b) using the standard NAEP estimates (presented in tables C4 and C5). There is no mean difference in trends, and although the differences varied somewhat from state to state, no differences were sufficient to change the result of a test for statistical significance.

Table 16. Mean difference in mathematics performance gains between 2000 and 2003, based on NAEP full population estimates (FPE) versus standard NAEP estimates (SNE), by grade

Level	Number of states	Mean difference in gain FPE-SNE: State assessment	Mean difference in gain FPE-SNE: NAEP	Standard deviation of difference in gains: FPE-SNE: State assessment	Standard deviation of difference in gains: FPE-SNE: NAEP	Number of statistically significant differences between NAEP and state assessment gains: FPE	Number of statistically significant differences between NAEP and state assessment gains: SNE
Grade 4	24	0.0	0.0	0.2	0.7	14	14
Grade 8	22	0.0	0.0	0.2	0.6	11	11

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

Finally, we compare the differences between full population estimates and standard NAEP estimates on gap comparisons. Table 17 shows the average differences in the achievement gaps (in 4th and 8th grade mathematics 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 17, we subtract the NAEP-state assessment differences based on standard NAEP estimates from the NAEP-state assessment differences based on full population estimates.

Table 17. Mean difference in gap measures of mathematics 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	Number of statistically significant differences between NAEP and state assessment gaps: SNE
	Black-White	25	0.8	0.8	14	17
Grade 4	Hispanic-White	14	-0.2	0.7	8	8
	Poverty	31	-0.2	0.6	12	12
	Black-White	20	0.2	0.5	7	9
Grade 8	Hispanic-White	14	0.0	0.5	6	5
	Poverty	28	-0.4	0.6	6	6

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 Mathematics Assessments: Full population estimates. The National Longitudinal School-level State Assessment Score Database (NLSLSASD) 2004.

Overall, the gap comparison results for standard NAEP estimates are similar to the results for full population estimates. Both sets of estimates agreed that NAEP measures of gaps were significantly larger than state assessment measures of those gaps in 49 comparisons and state assessments of gaps were larger in two comparisons. In only eight cases did the two sets of estimates disagree; and these disagreements did not exhibit any bias for one or the other method to estimate NAEP gaps to be relatively larger or smaller than state assessment gaps (four in each direction).

In summary, for measurement of gains in mathematics achievement since 2000 and minority and poverty gaps in mathematics achievement, as well as for correlations of percentages meeting mathematics standards between NAEP and state assessments, the choice to use full population estimates or standard NAEP estimates has only minor effects on the outcomes of comparisons between NAEP and state assessment results. That is, changes in exclusion rates between the 2000 and 2003 NAEP mathematics assessments and differences in exclusion rates between subpopulations had only minor effects on these NAEP-state assessment comparisons. That does not imply, it should be pointed out, that the use of these two different methods would yield the same results in comparisons of NAEP mathematics achievement gains and gaps between states, comparisons not undertaken in this report.

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 18. At grade 4, the median match rate across states was 99.1 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 over 99 percent indicates that the schools we missed tended to be schools carrying less weight in computing state averages from the NAEP

Table 18. Weighted and unweighted percentages of NAEP schools matched to state assessment records in mathematics, by grade and state: 2003

State/jurisdiction	Grade 4		Grade 8	
	unweighted	weighted	unweighted	weighted
Alabama	99.1	97.7	99.0	99.2
Alaska	100.0	100.0	100.0	100.0
Arizona	93.4	91.7	96.6	96.1
Arkansas	100.0	100.0	100.0	100.0
California	99.2	99.0	99.5	99.9
Colorado	96.8	96.0	98.2	98.3
Connecticut	100.0	100.0	100.0	100.0
Delaware	92.0	91.7	97.3	98.4
District of Columbia	87.3	89.1	73.7	83.4
Florida	98.1	98.3	99.0	99.1
Georgia	96.2	96.9	96.6	95.4
Hawaii	100.0	100.0	83.6	98.5
Idaho	100.0	100.0	100.0	100.0
Illinois	99.4	99.5	100.0	100.0
Indiana	100.0	100.0	100.0	100.0
Iowa	97.8	98.6	98.3	98.5
Kansas	99.3	98.3	99.2	99.6
Kentucky	100.0	100.0	100.0	100.0
Louisiana	100.0	100.0	100.0	100.0
Maine	98.7	99.9	99.1	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	99.9	100.0	100.0
Mississippi	99.1	99.0	100.0	100.0
Missouri	100.0	100.0	100.0	100.0
Montana	99.4	99.9	100.0	100.0
Nebraska	90.4	97.9	91.3	98.2
Nevada	98.2	96.3	97.0	96.4
New Hampshire	99.2	99.1	86.9	86.1
New Jersey	100.0	100.0	100.0	100.0
New Mexico	98.3	98.8	93.8	94.6
New York	98.0	98.3	97.3	98.2
North Carolina	98.7	99.6	98.5	98.4
North Dakota	97.6	99.6	100.0	100.0
Ohio	98.2	92.3	75.2	66.7
Oklahoma	100.0	100.0	99.2	99.5
Oregon	99.2	98.8	100.0	100.0
Pennsylvania	90.4	89.4	100.0	100.0
Rhode Island	99.1	99.6	98.1	98.4
South Carolina	97.2	98.2	99.0	98.7
South Dakota	89.8	98.4	89.1	98.8
Tennessee	100.0	100.0	99.1	96.5
Texas	99.0	97.8	98.6	95.9
Utah	98.3	96.9	100.0	100.0
Vermont	99.4	97.2	100.0	100.0
Virginia	95.7	93.8	96.3	94.0
Washington	99.1	99.7	100.0	100.0
West Virginia	100.0	100.0	100.0	100.0
Wisconsin	100.0	100.0	100.0	100.0
Wyoming	97.6	99.7	93.3	99.4

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

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 jurisdiction with a matching rate less than 90 percent were the District of Columbia (87 percent) and South Dakota (90 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 98.4 percent.

For grade 8, we were able to match more than 90 percent of the schools in all but five jurisdictions: District of Columbia (74 percent), Ohio (75 percent), Hawaii (84 percent), New Hampshire (87 percent), and South Dakota (89 percent). For Ohio, we do not include any grade 8 results in this report; and for Hawaii and South Dakota, the weighted match rates are very high. However, for the District of Columbia and New Hampshire, the lower match rate may offer one explanation for any discrepancies that are found between NAEP and state testing 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, more than 85 percent of the student population was represented in the analyses for all jurisdictions except 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 missing schools 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 19.

Across the states for which gap profiles are included in this report, the median percentages of Hispanic students and disadvantaged students included in grade 4 analyses is 85 percent, and the median percentage of Black students is 87 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, Missouri, and Wisconsin. Connecticut and Idaho Hispanic gap analyses are based on fewer than two-thirds of the Hispanic students in each state; and poverty gap analyses in Delaware, Missouri, New York, Vermont, and New Hampshire are based on fewer than two-thirds of the disadvantaged students in these states, based on NAEP estimates.

At grade 8, the situation is better, because with larger schools, fewer minority data are suppressed in state assessment files. The median percentages included in gap analyses are 94 percent for Blacks, 92 percent for Hispanics, and 90 percent for disadvantaged students; and there are no states in which analyses are based on fewer than 70 percent of the minority students in NAEP files.

Table 19. Percentages of NAEP student subpopulations in grades 4 and 8 included in comparison analysis in mathematics, 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	90.1	—	—	93.2	89.7
Alaska	—	—	—	—	—	—
Arizona	—	—	77.5	90.7	—	—
Arkansas	95.8	84.4	—	—	98.1	92.6
California	—	—	94.3	98.0	94.9	97.6
Colorado	—	—	—	—	—	—
Connecticut	61.1	74.3	61.1	75.2	79.2	83.3
Delaware	58.6	95.7	—	—	64.6	97.1
District of Columbia	—	—	—	—	71.9	76.5
Florida	95.3	96.8	88.4	97.7	95.8	98.2
Georgia	90.2	95.3	—	—	93.0	95.8
Hawaii	—	—	—	—	93.4	96.7
Idaho	—	—	63.7	84.8	—	—
Illinois	81.0	93.8	87.7	89.7	83.9	89.8
Indiana	80.7	94.8	—	—	92.1	99.2
Iowa	—	—	—	—	—	—
Kansas	57.5	—	—	—	83.9	86.5
Kentucky	85.2	—	—	—	91.6	100.0
Louisiana	95.4	98.8	—	—	98.8	96.8
Maine	—	—	—	—	—	—
Maryland	—	—	—	—	—	—
Massachusetts	69.5	—	80.2	89.7	—	—
Michigan	—	—	—	—	—	—
Minnesota	—	—	—	—	81.9	—
Mississippi	93.1	93.8	—	—	92.1	88.7
Missouri	62.9	85.9	—	—	63.9	74.5
Montana	—	—	—	—	—	—
Nebraska	—	—	—	—	—	—
Nevada	—	—	93.9	97.3	83.7	87.4
New Hampshire	—	—	—	—	59.6	—
New Jersey	85.3	91.3	85.6	92.8	87.0	97.0
New Mexico	—	—	72.2	73.8	71.6	85.9
New York	77.2	83.5	84.7	85.6	65.0	70.2
North Carolina	98.0	97.8	—	—	99.5	97.4
North Dakota	—	—	—	—	—	—
Ohio	87.6	—	—	—	81.4	—
Oklahoma	96.6	93.2	—	—	—	—
Oregon	—	—	—	93.5	—	—
Pennsylvania	78.4	94.9	—	—	81.0	96.9
Rhode Island	—	—	89.5	97.4	—	—
South Carolina	91.2	84.9	—	—	93.6	88.6
South Dakota	—	—	—	—	73.3	78.5
Tennessee	96.6	86.6	—	—	97.7	81.1
Texas	88.7	96.1	96.8	97.4	—	—
Utah	—	—	—	—	—	—
Vermont	—	—	—	—	61.1	73.0
Virginia	87.1	96.9	—	—	—	—
Washington	—	—	70.1	—	—	—
West Virginia	—	—	—	—	—	—
Wisconsin	54.5	—	—	—	86.7	87.2
Wyoming	—	—	—	—	95.4	92.8

— Not available.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Mathematics 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 20 and 21 show the percentages of students meeting the primary standard for NAEP samples and states' published reports of mathematics 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 20 and 21 than sampling variation.

Table 20. Percentages of grade 4 students meeting primary standard of mathematics achievement in NAEP samples and states' published reports, by state: 2000 and 2003

	NAEP		State reports	
State/jurisdiction	2000	2003	2000	2003
Alabama	—	—	—	—
Alaska	—	67.2	65.0	64.8
Arizona	40.6	49.8	35.0	49.0
Arkansas	35.6	60.1	37.0	60.0
California	52.3	—	51.0	—
Colorado	—	86	—	56.0
Connecticut	63.7	58.6	60.2	60.4
Delaware	—	72.7	62.0	71.0
District of Columbia	36.9	—	—	—
Florida	—	55.4	46.0	54.0
Georgia	61.6	73.9	62.0	74.0
Hawaii	64.8	67.4	—	—
Idaho	—	76.6	—	77.5
Illinois	52.7	68.3	57.3	68.3
Indiana	—	—	73.0	66.0
Iowa	—	75.2	71.0	75.0
Kansas	59.3	73.7	62.4	73.6
Kentucky	31.1	37.9	31.3	38.1
Louisiana	11.3	15.5	12.0	16.0
Maine	23.3	29.1	23.0	28.0
Maryland	45.6	—	—	55.0
Massachusetts	41	38	40.0	40.0
Michigan	76.9	—	74.8	65.0
Minnesota	47.9	58.2	45.6	57.0
Mississippi	—	74	—	73.7
Missouri	36.6	36.7	36.7	37.2
Montana	—	75.3	—	73.0
Nebraska	60	—	—	—
Nevada	—	50.8	—	—
New Hampshire	—	80.3	40.0	42.0
New Jersey	—	67.6	65.8	—
New Mexico	—	42	—	—
New York	67.8	78.8	65.0	79.0
North Carolina	84.6	92.2	—	92.1
North Dakota	—	59	—	—
Ohio	42.3	59	—	58.0
Oklahoma	85.8	69.3	85.0	72.0
Oregon	67.2	77.8	70.0	76.0
Pennsylvania	—	56.8	52.0	56.3
Rhode Island	20.7	41.8	—	42.6
South Carolina	22.9	32.6	24.0	33.7
South Dakota	—	72.5	—	—
Tennessee	—	—	—	—
Texas	88.5	—	87.0	87.0
Utah	—	—	—	—
Vermont	47	52.9	47.3	53.0
Virginia	—	—	63.0	—
Washington	—	54	41.8	55.2
West Virginia	—	—	—	—
Wisconsin	71.9	—	74.0	71.0
Wyoming	25.6	36.4	27.0	37.0

— 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), 2000 and 2003 Mathematics Assessments: Full population estimates. State reports are from state education agency websites.

Table 21. Percentages of grade 8 students meeting primary standard of mathematics achievement in NAEP samples and states' published reports, by state: 2000 and 2003

	NAEP		State report	
State/jurisdiction	2000	2003	2000	2003
Alabama	—	—	—	—
Alaska	—	65.2	39.0	63.8
Arizona	18.2	20.6	18.0	21.0
Arkansas	13.3	21.9	14.0	22.0
California	48	—	48.0	—
Colorado	—	68.2	35.0	38.0
Connecticut	57.2	56.0	54.8	56.1
Delaware	—	48.2	41.0	47.0
District of Columbia	8.6	—	—	—
Florida	—	54.1	51.0	56.0
Georgia	54.8	66.4	54.0	67.0
Hawaii	60.5	54.1	—	—
Idaho	—	52.5	—	53.0
Illinois	45.8	53.6	46.8	53.1
Indiana	—	—	63.0	66.0
Iowa	—	71.7	—	73.6
Kansas	55.3	59.3	54.6	60.0
Kentucky	26.0	31.9	25.2	30.9
Louisiana	7.3	8.7	8.0	8.0
Maine	20.6	17.1	21.0	18.0
Maryland	51.1	—	—	39.7
Massachusetts	33.7	38.1	34.0	37.0
Michigan	—	—	—	52.0
Minnesota	—	—	—	—
Mississippi	—	46.0	—	48.1
Missouri	12.9	13.4	14.0	13.9
Montana	—	70.3	—	69.0
Nebraska	59.7	—	—	—
Nevada	—	—	—	—
New Hampshire	—	—	—	—
New Jersey	—	56.2	59.7	—
New Mexico	—	39.4	—	—
New York	41.4	54.0	40.0	51.0
North Carolina	80.7	82.2	—	82.4
North Dakota	—	43.6	—	—
Ohio	—	—	—	—
Oklahoma	70.5	71.3	71.0	73.0
Oregon	54.7	57.6	56.0	59.0
Pennsylvania	—	51.5	52.0	51.3
Rhode Island	26.5	35.3	—	35.2
South Carolina	19.3	20.2	20.0	19.2
South Dakota	—	57.5	—	—
Tennessee	—	—	—	—
Texas	89.8	—	91.0	72.0
Utah	—	—	—	—
Vermont	45.6	51.7	47.0	51.7
Virginia	—	—	71.0	—
Washington	—	36.4	28.2	36.8
West Virginia	—	—	—	—
Wisconsin	42.9	—	30.0	73.0
Wyoming	31.5	35.1	32.0	35.0

— 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), 2000 and 2003 Mathematics Assessments: Full population estimates. State reports are from state education agency websites.

References

Campbell, J.R., Hombo, 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

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 reading assessment results: 2003* (NCES 2008-474). 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.