

# 5

## Comparing Achievement Gaps

A primary objective of federal involvement in education is to ensure equal opportunity for all students, including minority groups and those living in poverty (USDE 2002). NAEP has shown that although there have been gains since 1970, the average reading achievement of certain minority groups lags behind that of other students in both the elementary and secondary grades.<sup>32</sup> Numerous programs nationwide are aimed at reducing the reading achievement gap between Black and Hispanic students and White students, as well as between students in high-poverty and low-poverty schools; state assessments are monitoring achievement to determine whether, in their state, the gap is closing.

In compliance with *No Child Left Behind*, state education agencies now report school reading achievement results separately for minorities, for students eligible for free or reduced price lunch, for students with disabilities, and for English language learners (USDE 2002). These reports can be used to assess how successfully schools are narrowing the achievement gaps and to identify places needing assistance in narrowing their gaps.

Fair and unbiased measurement of the achievement of students from different cultural backgrounds is particularly difficult, and test developers try hard to remove test items that might unfairly challenge some groups more than others. In spite of these efforts, some state assessments may be more sensitive to achievement gaps and their narrowing than others. Comparison of NAEP measurement of reading achievement gaps to state assessment results can shed light on such differences.

The main objective of this part of the report is to compare the measurement of reading achievement gaps by state assessments and NAEP. Specifically, we compare three types of gaps:

- the Black-White achievement gap
- the Hispanic-White achievement gap
- the poverty gap: achievement of students qualifying for free or reduced-price lunch (i.e., disadvantaged students) versus those who do not qualify.<sup>33</sup>

---

32. Campbell, Hombo, and Mazzeo (2000)

33. We refer to students eligible for free/reduced price lunch as (economically) disadvantaged students.

The focus of these comparisons is not on differences in gaps between states but on differences between NAEP and state measurement of the gap in the same schools in the same state.

## POPULATION PROFILES OF ACHIEVEMENT GAPS

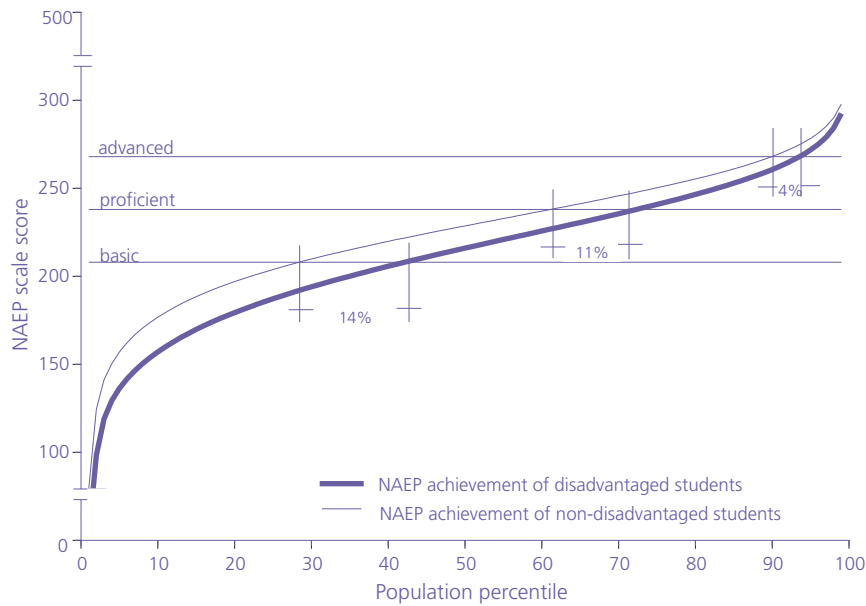
Achievement gaps for whole subpopulations, such as Black students, Hispanic students, or economically disadvantaged students, are complex. What causes one segment of a disadvantaged population to achieve at a lower level may be quite different from the barriers faced by another segment of the same subpopulation. It is easy to forget that in the context of a population achievement gap, there are still many students in the disadvantaged group who achieve at a higher level than typical for non-disadvantaged groups. Expressing a reading achievement gap as a single number (the difference in the percentages of children in two groups who meet a standard) hides a great deal of information about the nature of gaps.

Moreover, as Paul Holland (2002) has shown, an achievement gap is also likely to mislead readers because of the differential placement of the standard relative to the distribution of achievement in the two populations. Figure 19 below is shown as an example to illustrate how the grade 4 poverty gap in reading achievement in 2003, which is about 13 points on the NAEP scale at the median, is larger in the lower part of the achievement distribution than in the higher part of the distribution. As a result, the gap is 14 percent in achieving the basic level (the distance between the points at which the graphs cross the basic criterion of 208 on the NAEP scale), 11 percent in achieving the proficient level, and four percent in achieving the advanced level. The graph, or population profile, conveys significantly more information about the poverty gap in reading than does a simple comparison of the percentages achieving the standards.

Holland points out that this effect is more striking when examining reduction in gaps. If we hypothetically suppose that at some future date all students would gain 20 points on the NAEP reading achievement scale, the population profiles would appear as in figure 20.

In this case, the gaps in percent proficient, 14 percent, and percent advanced, 9 percent, would be larger than in 2003, while the gap in percent achieving the basic level, 11 percent, would be smaller. Even though the achievement gaps remain constant, they appear to increase or decrease, depending on the relationship between the standards and the distribution of student achievement. Even though the gap in the percentage of students achieving the basic level might be reduced from 14 percent to 11 percent, the gap in reading skills at that level would be just as large as before, larger than among higher-achieving segments of the disadvantaged and non-disadvantaged populations. Even though the gap in the percentage of students meeting the advanced standard might increase from 4 percent to 9 percent, it would be inappropriate to conclude that educators were allowing the gap among the highest achievers in the two subpopulations to increase.

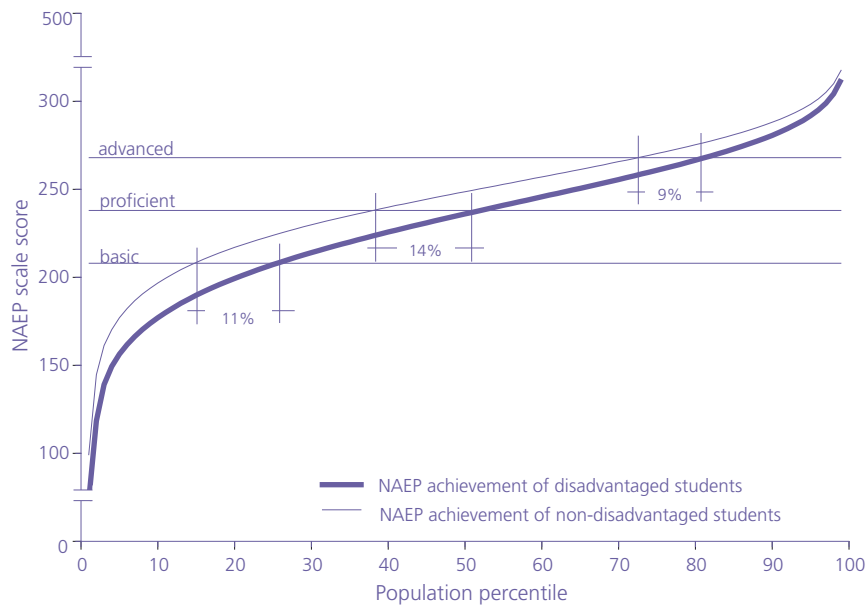
Figure 19. Population profile of the NAEP poverty gap in grade 4 reading achievement: 2003



NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged.

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.

Figure 20. NAEP poverty gap from a hypothetical uniform 20-point increase on the grade 4 reading achievement scale



NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged.

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.

If gap measurement must be carried out in terms of percentages meeting standards, it is essential not to be misled by comparisons between gaps measured at different points on the achievement continuum. This effect makes it clear that comparison of NAEP and state assessment measurement of gaps in terms of percentages meeting standards must refer to the same standard for both assessments. Therefore, because individual scale values are not available for state assessments, we must measure gaps at the level of each state's primary standard.

It is important to note, however, that measuring the gap at each state's particular standard renders comparisons of gaps between states uninterpretable, because the standards are different in each state. And, even though NAEP applies the same set of standards in all states to produce the biennial Nation's Report Card, comparisons of achievement gaps in different states must be interpreted in the context of variations in the position of the NAEP standards with respect to states' achievement distributions. Comparisons of gaps between states can be found in the Nation's Report Card (<http://nces.ed.gov/nationsreportcard>).

There are three major limitations in the data that further affect the interpretation of comparisons of gaps as measured by NAEP versus state assessments. The first limitation is that the state assessment data are only available at the school and grade level for the various population groups, not for individuals. The second is that percentages of subgroups meeting standards are suppressed in many schools, due to small samples. The third is that the separate scores for subpopulations are not available in the NLSLSASD before 2002, limiting the possibility of reporting reduction in gaps over time.

The state assessment data's limitation to school-level aggregate percentages of subgroups achieving standards means that each student is represented as the average of that student's population group in a school. As a result, variability in performance within each group in a school is not captured. The variability across schools using group averages is lower than the variability across schools using individual student scores. Unfortunately, to compare NAEP and state assessment gaps with each other, we must also limit the NAEP data to school averages for subgroup.

In addition, school-level scores are subject to *suppression* when the number of students tested in a subgroup is so small that the state prohibits the release of scores. Suppression rules vary between states; typically, scores are suppressed when fewer than 5 or 10 students are included in the average score. To avoid unreliable NAEP results in gap comparisons between NAEP and state assessment reports, we have omitted from analysis any schools with fewer than three tested subgroup members.<sup>34</sup>

34. Including percentages based on one or two students overestimates the frequency of observing extreme percentages: with one student, the percentage is either 0 or 100. Because small schools in the NAEP sample may be weighted to represent large numbers of small schools, this distorts some population profiles by overestimating the percentages of students in the extreme categories of 0 percent achieving the standard and 100 percent achieving the standard. Suppressing the cases based on one or two students more closely matches the state assessment practices of suppressing scores based on small sample sizes.

## READING ACHIEVEMENT GAPS IN 2003

The State Profile section of this report (appendix D) displays three types of achievement gaps similar to Figure 19, for all states with available data: the Black-White achievement gap, the Hispanic-White achievement gap, and the poverty achievement gap. These are introduced in the following pages through the presentation of population profiles of gaps showing the aggregation of percentages of students meeting states' primary standards. The graphs are not intended to reflect national average achievement gaps because they represent only some states and some schools, but they are informative. Although the graphs portray the aggregate achievement gap as measured against different standards across states, the general size of the gaps, in terms of standards in place in the nation in 2003, is apparent.

An aggregate population profile of the poverty gap in grade 4 reading achievement for the states included in this report is shown in the following series of four graphs, which compare the reading achievement of disadvantaged students (defined as students eligible for free/reduced price lunch) with other students. Figures 21 and 22 display the differential achievement, first as measured by NAEP (in figure 21) and second as measured by state assessments (in figure 22).<sup>35</sup>

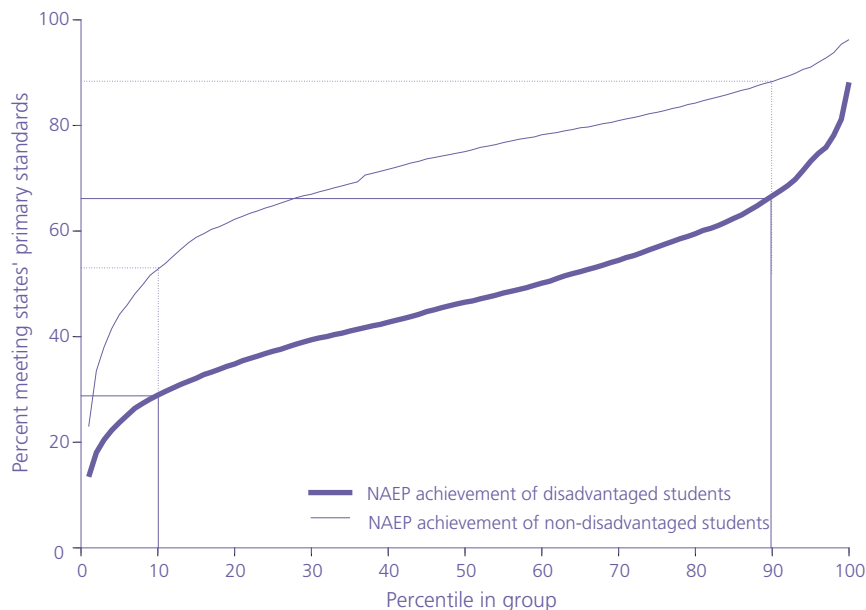
In figure 21, the vertical (y) axis measures reading achievement. Due to limitations on the data available on state assessment results, reading achievement cannot be graphed for each individual student. Instead, it is measured by the percent of students in a school meeting the state's primary reading achievement standard. That is, for each student in a subgroup (such as disadvantaged students), the reading achievement measure is the percent of students in his or her subgroup in his or her school meeting the standard. Thus, within a particular school, all members of the subgroup have the same reading achievement measure, which is the average of all their individual achievement scores. The horizontal (x) axis represents all the students in the subgroup in the state or nation, arrayed from those with the lowest reading achievement measure on the left to the highest reading achievement measure on the right. The units on the horizontal axis are percentiles, from 0 to 100; that is, percentages of the student's subgroup with equal or lower reading achievement measures. Figure 21 serves a dual purpose. First, it arrays the disadvantaged student population (shown by the darker line) by percentile, from those in the schools where the fewest disadvantaged students meet the standard to those in schools where the

35. A school at the 50th percentile on a population profile has an average performance that is higher than the average performance in schools serving half the students in the state and lower than the average performance in the schools serving the other half of the population (except for those in the school itself, of course). However, there may be different numbers of schools serving the upper and lower halves of the population. For example, for minority population profiles, there may be 500 (small) schools serving students in the upper half of the (minority) student population and 100 (larger) schools serving students in the lower half of that population, so the population percentile is not literally a percentile of schools. It is a percentile of the student population served by the schools.

most disadvantaged students meet the standard. Second, it also arrays the non-disadvantaged student population (shown by the lighter line) by percentile. Thus, two population profiles can be superimposed on the same graph to display the achievement gap between them.

The population profiles in figure 21 can be read as follows. Focus first on comparing the highest achievers among disadvantaged students versus the highest achievers among non-disadvantaged students. Consider the 90th percentile of each population as an example. The vertical line at the 90th percentile crosses the *disadvantaged* line at 67 percent meeting the standard. That means that 90 percent of the population of disadvantaged students are attending schools where fewer than 67 percent of the disadvantaged students meet the standard and the other 10 percent are in schools where more than 67 percent of the disadvantaged students meet the standard. In other words, students at the 90th percentile of the disadvantaged population are attending schools where 67 percent of the disadvantaged students meet the standard.

Figure 21. School percentage of economically disadvantaged and non-disadvantaged students meeting states' primary grade 4 reading standards as measured by NAEP, by percentile of students in each subgroup: 2003



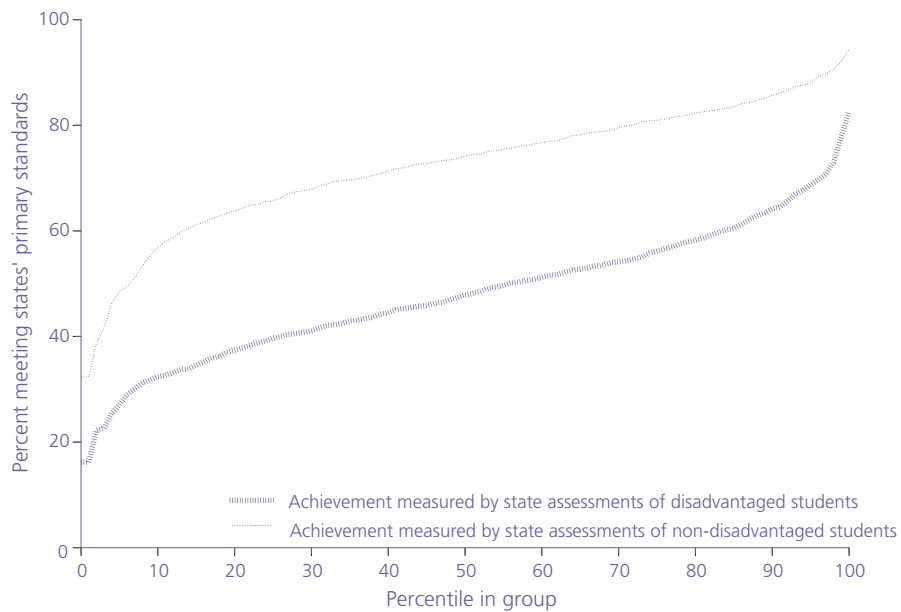
NOTE: Primary standard is the state's standard for *proficient* performance. Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.



By comparison, the 90th percentile crosses the *non-disadvantaged* line in figure 21 at 88 percent meeting the standard. That means that 90 percent of *non-disadvantaged students* are attending schools where fewer than 88 percent of *non-disadvantaged students* meet the standard and the other 10 percent are in schools where more than 88 percent of *non-disadvantaged students* meet the standard. We can say that the *students at the 90th percentile of the non-disadvantaged population are attending schools where 88 percent of the non-disadvantaged students meet the standard*. Thus, comparing 90th percentile for the disadvantaged student population versus the 90th percentile for the non-disadvantaged student population, there is a gap of 21 points (21=88-67) in percentages between the groups in their school meeting the standard.

**Figure 22. School percentage of economically disadvantaged and non-disadvantaged students meeting states' primary grade 4 reading standards as measured by state assessments, by percentile of students in each subgroup: 2003**



NOTE: Primary standard is the state's standard for *proficient* performance. Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

The graphs in figures 21 and 22 are aggregated across the states and schools for which subgroup percentages achieving reading standards are available.<sup>36</sup> Because the

36. States with fewer than 10 percent disadvantaged students or fewer than 10 NAEP schools with non-suppressed percentages for disadvantaged students are excluded due to unstable estimates.



primary standards vary from state to state, it is essential in comparing NAEP and state assessment results that the NAEP results are measured relative to each state's standard in that state. Corresponding population profiles of gaps for individual states are included in the State Profile section of this report (appendix D)—these aggregate profiles may provide context for interpreting individual state achievement gaps.

Figures 21 and 22 display similar pictures—throughout the middle range of the student populations, there is a fairly uniform gap of slightly more than 20 percentage points, and this gap is noticeably smaller in the extreme high and low percentiles. Nevertheless, some disadvantaged students attend schools where the percentages of disadvantaged students achieving the standard are greater than the percentages of non-disadvantaged students achieving the standard in other schools. For example, Figure 21 showed that 90th-percentile disadvantaged students are in schools where 67 percent of them meet the standards, which is a greater percentage than among the lowest quarter of the non-disadvantaged student population.<sup>37</sup>

All of the graphs are based on the NAEP schools, weighted to represent the population of fourth graders in each state. Because the use of the aggregate school-level percentages may have an effect on the position and shape of the population profile graphs, school-level percentages are presented in both NAEP and state assessment graphs. Appendix A presents a description of the method for constructing population profiles based on school-level aggregate achievement measures.

Figure 23 combines the NAEP and state assessment profiles for disadvantaged and non-disadvantaged students. The similarity of the NAEP and state assessment results in this figure is notable. Although some discrepancies are worth noting, the overall picture suggests that as a summary across 38 states, NAEP and state assessments are measuring the same poverty gap.

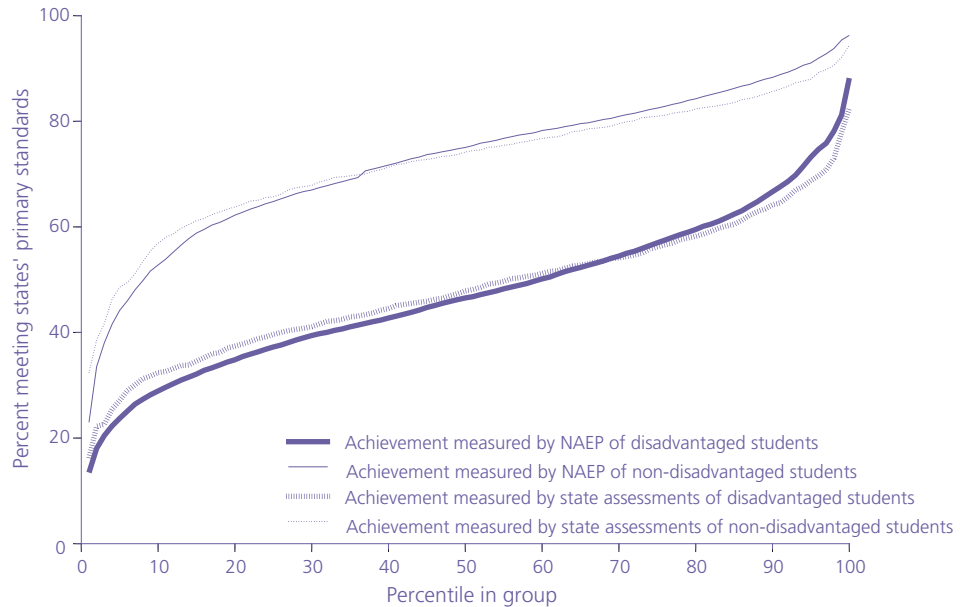
Compared to the average state assessment results, the NAEP population profiles appear to exhibit greater variation between the top and bottom of the distributions (the NAEP lines are above the state lines at the top of the distribution and below the state lines at the bottom of the distribution, i.e., the lines cross). In other words, there is greater variation in school-level achievement of reading standards measured by NAEP, within both the disadvantaged and non-disadvantaged populations, than in achievement measured by state assessments (Figure 23). Whether this is a real phenomenon or an artifact of the differences in design between NAEP and state assessments is not clear at this time and requires further study. One possibility is that because NAEP percentages meeting standards are generally based on fewer students in

37. Readers should not be confused by the use of *percent* and *percentile* for the two axes in the population profile graphs. These are two completely different measures, which happen to have similar names. For percentiles, there must be a person at the lowest percentile and another at the highest percentile, by definition, because the percentiles just rank the people from lowest (zero, or one, in some definitions) to highest (100). The percent achieving a standard, on the other hand, can be zero for everybody (a very high standard indeed!) or 100 for everybody, or anywhere in between. The only built-in constraint is that the graphs must rise from left to right - higher achieving segments of the (sub)population are ranked in higher percentiles of the (sub)population by definition.



each school than state assessment percentages, variation of NAEP school means may naturally be larger than variation of state assessment school means.

**Figure 23. School percentage of economically disadvantaged and non-disadvantaged students meeting states' primary grade 4 reading standards as measured by NAEP and state assessments, by percentile of students in each subgroup: 2003**



NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

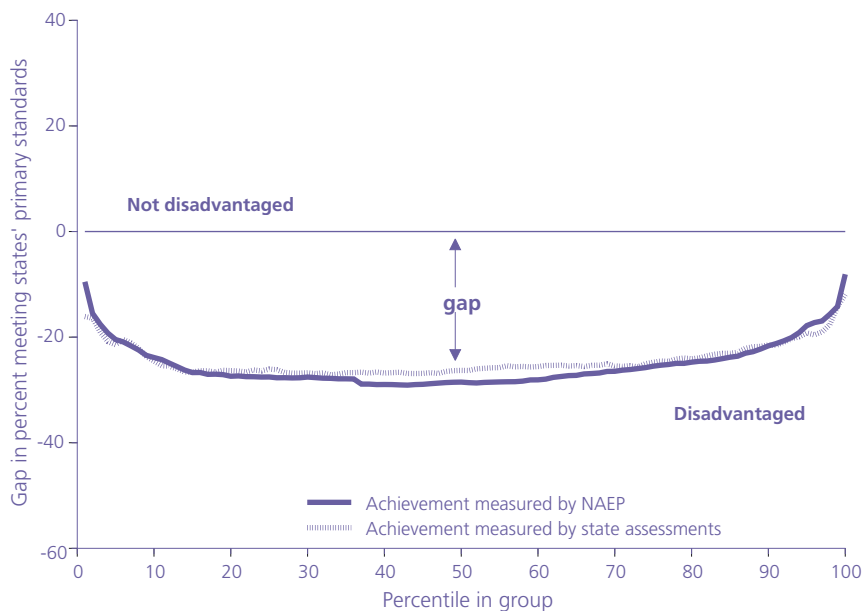
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

In figure 23, the size of the difference between NAEP and state assessment measurement of the poverty gap is difficult to separate visually from the size of the gap itself. In particular, in the highest achievement percentiles, NAEP reports higher achievement by both groups than the aggregate state assessments do. To focus on the differences between NAEP and state assessments, we eliminate the distracting information by graphing only the *difference* between the profiles (the achievement of the disadvantaged group, minus the achievement of the non-disadvantaged group). The result is the gap profile in figure 24, in which a *zero* gap is the goal, and current gaps fall below that goal. Gaps measured by NAEP and by state assessments are both shown.

In figure 24 it becomes clearer how similar the NAEP and state assessment measurement of poverty gaps are, overall. Although the gap profiles are very similar, NAEP is measuring a larger gap in the lower middle of the distribution than are state

assessments. From the 35th to the 55th percentile, NAEP measures a gap of 27 to 29 percent in achieving state standards, while state assessments average a 26 to 27 percent gap. On the other hand, below the 20th percentile and above the 70th percentile, the NAEP and average state assessment poverty gaps are virtually identical. Note that for individual state gap profiles (in appendix D) results of statistical significance tests are reported. Because the graphs presented here are not intended for inference about national patterns, no significance tests are reported.

**Figure 24. Profile of the poverty gap in school percentage of students meeting states' grade 4 reading achievement standards, by percentile of students in each subgroup: 2003**

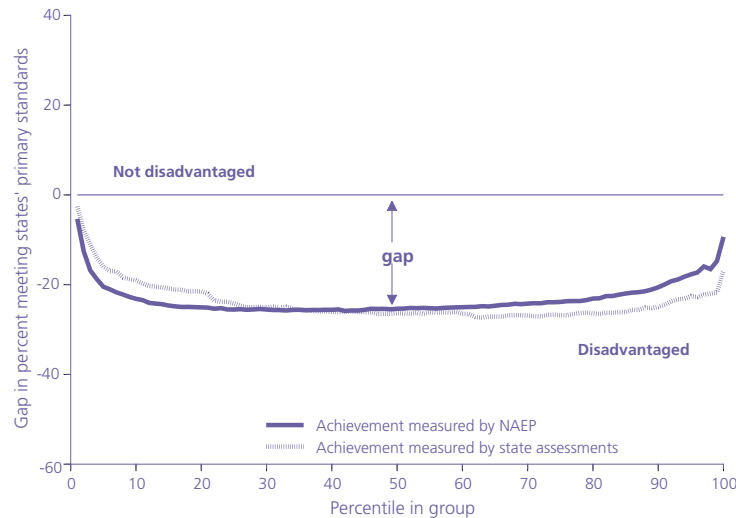


NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

Similar population profiles can be constructed for other comparisons between subpopulations. We present these to provide a general context in which to view the gaps for individual states displayed in appendix D. Figure 25 displays the same poverty gap information as is displayed in figure 24, but for grade 8. At grade 8, although the average gaps measured by NAEP and state assessments appear very similar, there is also a noticeable pattern, in that NAEP measures a slightly larger gap between lower-achieving disadvantaged and non-disadvantaged students and a slightly smaller gap between the higher-achieving members of the each group.

Figure 25. Profile of the poverty gap in school percentage of students meeting states' grade 8 reading achievement standards, by percentile of students in each subgroup: 2003



NOTE: Primary standard is the state's standard for *proficient* performance. Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

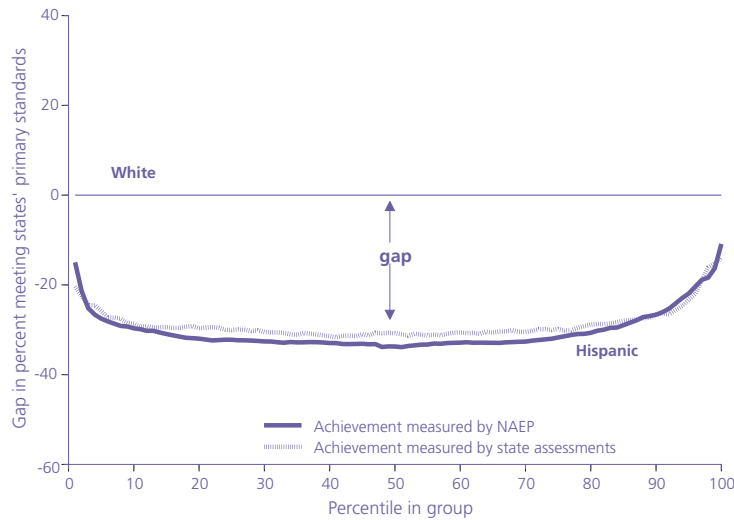
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment: Full population estimates. The National Longitudinal School-Level State Assessment Score Database (NLSLSASD) 2004.

It should be noted that this comparison, like all comparisons between NAEP and state assessment results in this report, is based on NAEP and state assessment results in the same set of schools. For example, if the state reported a percentage meeting their standard for disadvantaged students at a school, but the NAEP student sample in that school included no disadvantaged students, that school would not be included in the population profile of disadvantaged students. (Of course, that school might be included in the non-disadvantaged student profiles: the achievement gaps being reported here combine both between-school gaps and within-school gaps.)

Figures 26 and 27 provide aggregate population gap profiles for the Hispanic-White gap and the Black-White gap in grade 4 reading achievement.<sup>38</sup> The Hispanic-White gap is about 30 percent across much of the distribution, whether measured by NAEP or state assessments, while the Black-White gap pattern is similar to that for disadvantaged students—NAEP finds a slightly larger gap between the lower-achieving halves of the two populations. This effect may be related to what NAEP is

38. The aggregate Hispanic-White gap is based on results in 37 states, although sufficient sample sizes are available for comparison of results for individual states in only 14 states. The aggregate Black-White gap is based on results in 42 states, although sufficient sample sizes are available for comparison of results for individual states in only 26 states. The aggregate poverty gap is based on results in 38 states, although sufficient sample sizes are available for comparison of results for individual states in only 31 states.

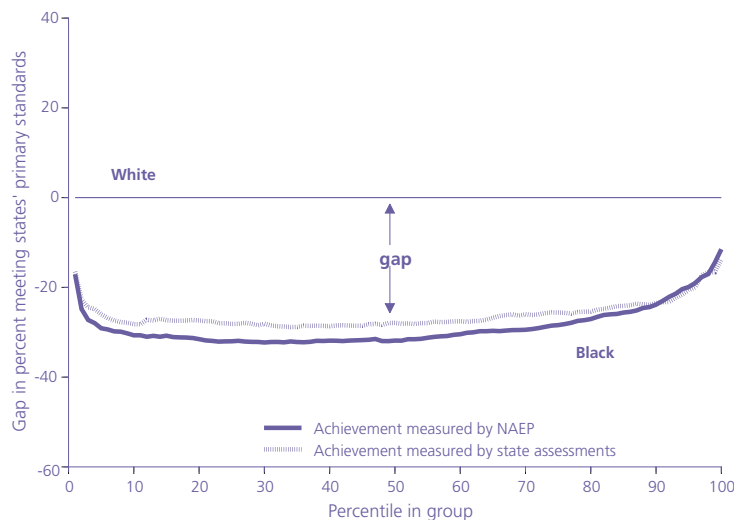
Figure 26. Profile of the Hispanic-White gap in school percentage of students meeting states' grade 4 reading achievement standards, by percentile of students in each subgroup: 2003



NOTE: Primary standard is the state's standard for *proficient* performance. Percentile in group refers to the percentage of the Hispanic or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

Figure 27. Profile of the Black-White gap in school percentage of students meeting states' grade 4 reading achievement standards, by percentile of students in each subgroup: 2003



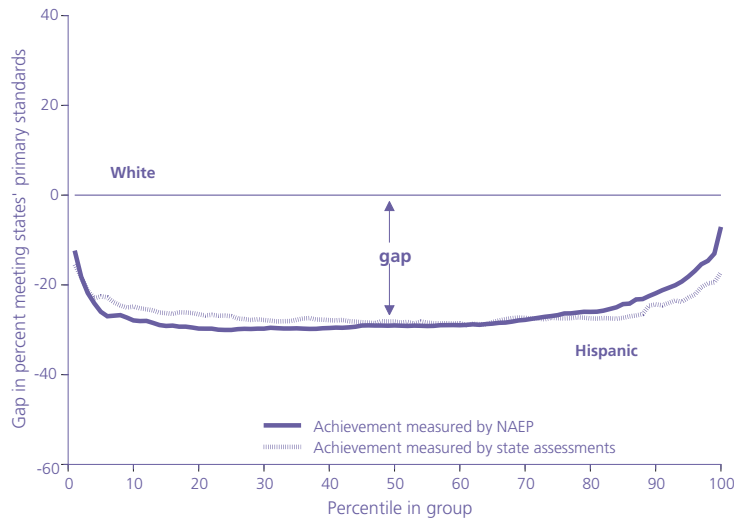
NOTE: Primary standard is the state's standard for *proficient* performance. Percentile in group refers to the percentage of the Black or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

measuring or it may be related to the fact that NAEP results are based on fewer students in each school than state assessment results are.

Corresponding aggregate grade 8 Hispanic-White and Black-White reading achievement gap profiles are shown in figures 28 and 29. The similarities and differences between grade 4 and grade 8 profiles we saw for the poverty gap are also found for these gap profiles: grade 8 gaps are similar overall to grade 4 gaps, but there is a tendency for NAEP to measure slightly smaller gaps in the higher-achieving parts of the populations and slightly larger gaps in the lower-achieving parts of the populations than the state assessments.

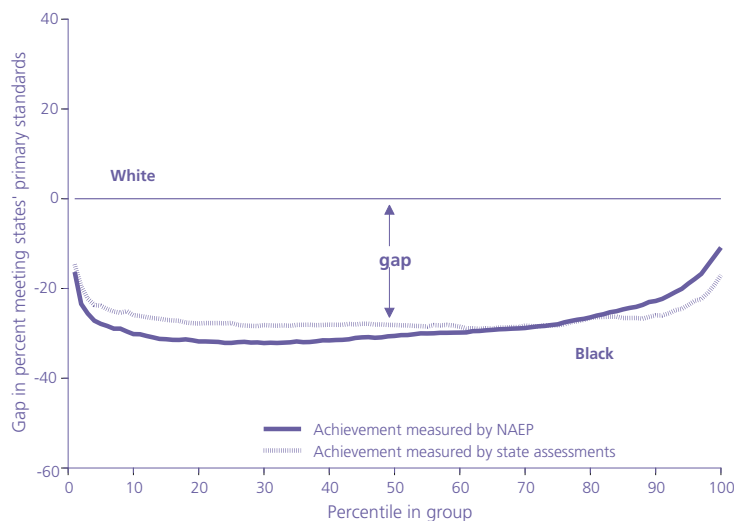
**Figure 28. Profile of the Hispanic-White gap in school percentage of students meeting states' grade 8 reading achievement standards, by percentile of students in each subgroup: 2003**



NOTE: Primary standard is the state's standard for *proficient* performance. Percentile in group refers to the percentage of the Hispanic or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

**Figure 29. Profile of the Black-White gap in school percentage of students meeting states' grade 8 reading achievement standards, by percentile of students in each subgroup: 2003**



NOTE: Primary standard is the state's standard for *proficient* performance. Percentile in group refers to the percentage of the Black or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.



Although the aggregate gap profile can identify small reliable differences in gaps as measured by NAEP and state assessments, the samples in individual states are not sufficiently large to detect small differences. The differences between the NAEP and state assessments in mean state gaps in percentage meeting the primary grade 4 reading standard are shown in table 10 for the states in which sufficient subgroup data are available. Although these statistics are based on the overall average, readers can examine Student's *t* test results for various parts of the distributions (halves and quartiles) for individual states in appendix D.

At grade 4, NAEP measured significantly larger mean gaps in 2003 than the state assessment did in 3 of 26 Black-White comparisons (Arkansas, Indiana, and New York); 1 of 14 Hispanic-White comparisons (Arizona); and 1 of 31 poverty comparisons (Indiana). In Delaware, on the other hand, the state assessment found a larger mean Black-White gap than NAEP did (although this may be affected by the fact that the schools available for comparison only represented 57 percent of the Black students in the state). In Nevada, the state assessment found a larger Hispanic-White gap.

At grade 8, the pattern is similar (Table 11). NAEP found significantly larger Black-White gaps in 2 of 20 states (Illinois and New York); significantly larger Hispanic-White gaps in 2 of 13 states (Illinois and Rhode Island); and significantly larger poverty gaps in 4 of 28 jurisdictions (District of Columbia, Hawaii, Illinois, and Vermont). By contrast, the state assessments found a larger Hispanic-White gap than NAEP did in Idaho; and larger poverty gaps than NAEP did in Arkansas, California, Florida, Nevada, and South Carolina.

Not very much should be made of these significant results until additional studies are performed. Examination of the individual state gap profiles in appendix D supports the conclusion that for the most part, NAEP and state assessments are measuring nearly the same gaps. For example, there are no states in which we were able to carry out comparisons and in which the NAEP gap is twice as large as the state assessment gap. Various factors, both substantive and methodological, may explain the tendency for NAEP to find slightly larger gaps.<sup>39</sup> These must be factors that differentially affect the performance of students in different groups.

Among such possible factors, on the methodological side, there could be differences in student motivation, in methods of analyzing the test scores, or in prevalence of testing accommodations. Similarly, on the substantive side, it is possible that variation in scores on a state assessment, which focuses on what is taught in the schools, is somewhat less related to cultural differences that children bring to their schoolwork, compared to NAEP, because NAEP aims for an overall assessment of reading achievement, including both culturally and school-related components of that performance.

---

39. The determination of those factors is beyond the scope of the report.



Table 10. Differences between NAEP and state assessments of grade 4 reading achievement race and poverty gaps, by state: 2003

State/ jurisdiction	Black-White	Hispanic-White	Poverty
Alabama	0.1	—	0.6
Alaska	—	—	—
Arizona	—	-7.7 *	—
Arkansas	-8.8 *	—	2.8
California	—	-0.3	2.4
Colorado	—	—	—
Connecticut	-2.0	3.1	1.2
Delaware	3.9 *	—	-0.0
District of Columbia	—	—	-0.3
Florida	-1.4	-0.7	3.3
Georgia	-4.3	—	-3.8
Hawaii	—	—	-3.4
Idaho	—	1.2	—
Illinois	-0.4	-2.5	-3.1
Indiana	-11.1 *	—	-7.3 *
Iowa	—	—	—
Kansas	-3.1	—	-8.7
Kentucky	0.6	—	-4.2
Louisiana	-3.5	—	-2.5
Maine	—	—	—
Maryland	—	—	—
Massachusetts	-4.3	-4.2	—
Michigan	—	—	—
Minnesota	—	—	1.2
Mississippi	-1.7	—	-3.0
Missouri	-2.9	—	—
Montana	—	—	—
Nebraska	—	—	—
Nevada	3.3	7.5 *	2.8
New Hampshire	—	—	-5.0
New Jersey	-3.6	-3.7	-3.3
New Mexico	—	-0.3	0.1
New York	-10.3 *	-6.1	-2.6
North Carolina	-3.0	—	-1.9
North Dakota	—	—	—
Ohio	-5.3	—	-4.4
Oklahoma	7.5	—	—
Oregon	—	—	—
Pennsylvania	-0.5	—	-2.6
Rhode Island	—	-4.7	—
South Carolina	-3.5	—	-1.3
South Dakota	—	—	-1.6
Tennessee	-0.9	—	2.0
Texas	1.4	-1.0	—
Utah	—	—	—
Vermont	—	—	-5.4
Virginia	-4.0	—	—
Washington	—	-1.1	—
West Virginia	—	—	1.2
Wisconsin	-4.2	—	-4.2
Wyoming	—	—	0.1

— Not available.

\* The NAEP-state difference is statistically significant at  $p < .05$ .

NOTE: A positive entry indicates that the state assessment reports the gap as larger than NAEP does; a negative entry indicates the state assessment reports the gap as smaller.

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.

Table 11. Differences between NAEP and state assessments of grade 8 reading achievement race and poverty gaps, by state: 2003

State/ jurisdiction	Black-White	Hispanic-White	Poverty
Alabama	-0.8	—	2.5
Alaska	—	—	—
Arizona	—	-4.2	—
Arkansas	-1.7	—	6.7 *
California	—	4.3	6.9 *
Colorado	—	—	—
Connecticut	4.2	6.1	4.5
Delaware	-2.8	—	-0.8
District of Columbia	—	—	-4.8 *
Florida	-0.1	1.0	5.6 *
Georgia	-3.1	—	-1.3
Hawaii	—	—	-3.8 *
Idaho	—	8.4 *	—
Illinois	-7.2 *	-5.8 *	-7.3 *
Indiana	0.3	—	-1.6
Iowa	—	—	—
Kansas	—	—	-0.5
Kentucky	—	—	3.1
Louisiana	0.5	—	1.6
Maine	—	—	—
Maryland	—	—	—
Massachusetts	—	—	—
Michigan	—	—	—
Minnesota	—	—	—
Mississippi	0.1	—	1.1
Missouri	-0.4	—	—
Montana	—	—	—
Nebraska	—	—	—
Nevada	-0.5	1.9	5.3 *
New Hampshire	—	—	—
New Jersey	4.5	-0.5	2.2
New Mexico	—	-3.0	-1.3
New York	-9.9 *	-5.0	-3.4
North Carolina	-0.2	—	1.9
North Dakota	—	—	—
Ohio	—	—	—
Oklahoma	—	—	—
Oregon	—	6.0	—
Pennsylvania	-1.3	—	-1.6
Rhode Island	—	-6.5 *	—
South Carolina	0.5	—	4.4 *
South Dakota	—	—	2.0
Tennessee	-1.2	—	4.0
Texas	-0.3	-3.8	—
Utah	—	—	—
Vermont	—	—	-6.3 *
Virginia	-4.8	—	—
Washington	—	—	—
West Virginia	—	—	2.7
Wisconsin	—	—	-1.8
Wyoming	—	—	1.7

— Not available.

\* The NAEP-state difference is statistically significant at  $p < .05$ .

NOTE: A positive entry indicates that the state assessment reports the gap as larger than NAEP does; a negative entry indicates the state assessment reports the gap as smaller.

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.

## ACHIEVEMENT GAP REDUCTION BETWEEN 2002 AND 2003

Culturally related gaps in reading achievement have persisted over the past 30 years in NAEP surveys of achievement in reading (NCES 2005). Even though educational policymakers and educators strive to find ways to reduce these gaps, they continue to exist. It is therefore of interest to know whether NAEP and state assessments are reporting the same measure of progress in reducing gaps.

Estimates based on NAEP are limited to the years in which NAEP has been administered according to a design that permits separate state reporting. For reading, these years were 1992, 1994, 1998, 2002, and 2003. On the other hand, the availability of school-level scores disaggregated by subgroup in the NLSLSASD is limited to 2002 and 2003, although such data may continue to be added in the future. Therefore, for this report, the only measures of trends in gap reduction are for the one-year period between the 2001-02 and 2002-03 school years.

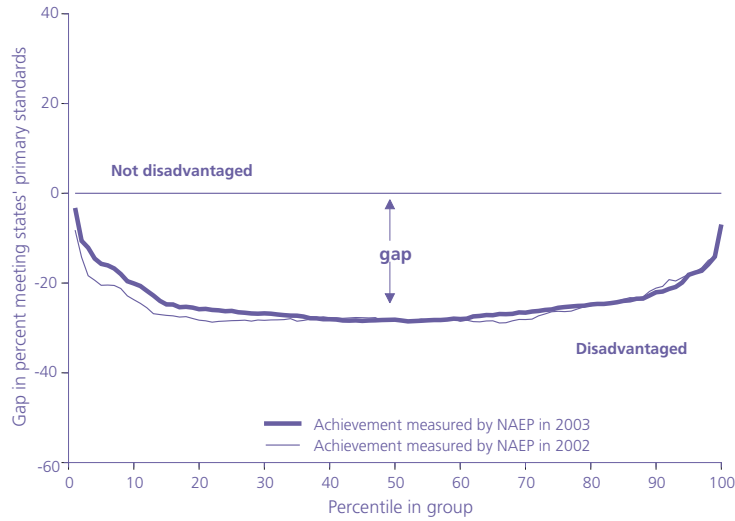
To compare NAEP and state assessment measurement of changes in achievement gaps from 2002 to 2003, we computed gap profiles like those in figures 24 through 29 for each year. The simple difference, by subtraction, between the gap profiles for two years yields a profile of the gap reduction. For illustration, figures 30 and 31 display the poverty gap profiles in 2002 and 2003, aggregated across the 18 states with available scores for both years, as measured by NAEP and state assessments, respectively. From these profiles, it is clear that there was not a great deal of change in the poverty gap in that year. However, in the lower-achieving halves of the two populations, the state assessments in aggregate appear to have recorded somewhat larger poverty gaps in 2003 than in 2002. This is not seen in the NAEP measurement.

The gap trends can be compared more directly by subtracting the 2003 gap from the 2002 gap.<sup>40</sup> If the gap reduced, then this subtraction yields a positive result; conversely, if the gap widened, the subtraction yields a negative number. These trend comparisons, as measured by NAEP and state assessments, are displayed in figures 32 (for grade 4) and 33 (for grade 8). The primary message in these figures is that in the single year from 2002 to 2003 there was no reliable pattern of reduction of the poverty gap across the 17 states (grade 8) or 18 states (grade 4) for which data were available for both years.

Hispanic-White and Black-White trend comparisons, analogous to the poverty trend comparisons in figures 32 and 33, are displayed in figures 34 through 37. Again, there appears to be no systematic pattern of improvement with the passage of a single year, but these results are based on only 20 states, and on only the schools in those states for which there were sufficient numbers of minority students for the state score breakdowns to be released.

40. The achievement gaps, or deficits, are displayed as negative numbers in this report, subtracting the achievement of the comparison group from the achievement of the focal group.

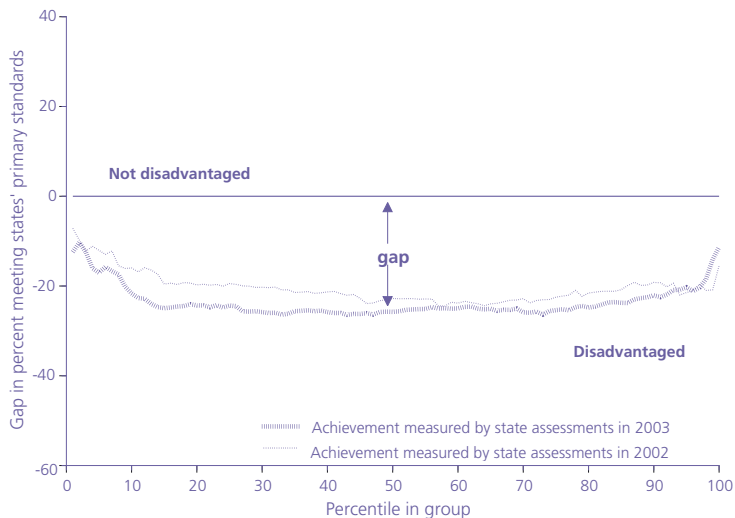
Figure 30. Profile of the poverty gap in school percentage of students meeting states' grade 4 reading achievement standards as measured by NAEP, by percentile of students in each subgroup: 2002 and 2003



NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

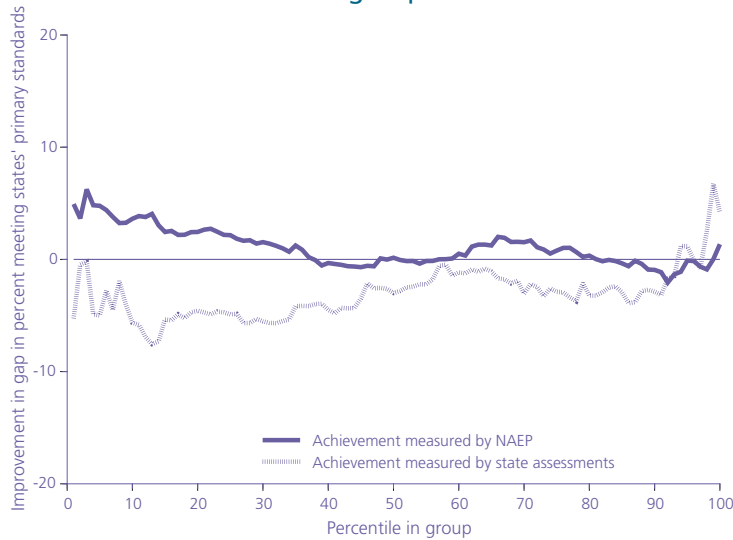
Figure 31. Profile of the poverty gap in school percentage of students meeting states' grade 4 reading standards as measured by state assessments, by percentile of students in each subgroup: 2002 and 2003



NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

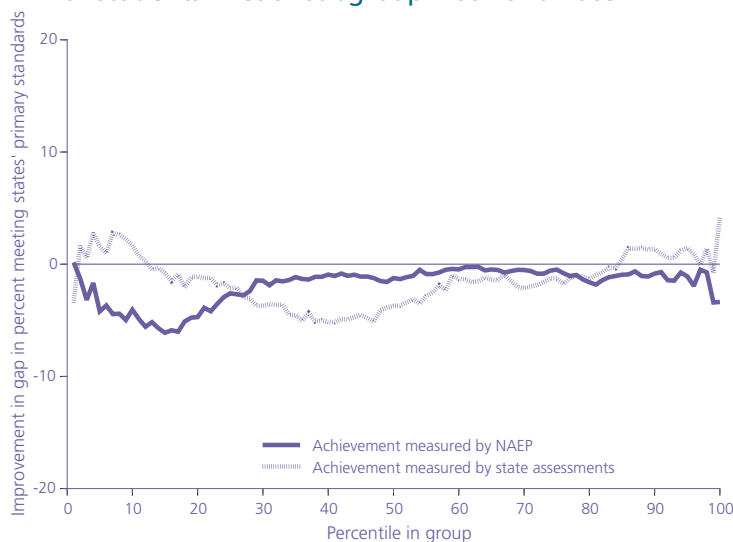
**Figure 32. Profile of the reduction of the poverty gap in school percentage of students meeting states' primary grade 4 reading achievement standards, as measured by NAEP and state assessments, by percentile of students in each subgroup: 2002 and 2003**



NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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 33. Profile of the reduction of the poverty gap in school percentage of students meeting states' primary grade 8 reading achievement standards, as measured by NAEP and state assessments, by percentile of students in each subgroup: 2002 and 2003**

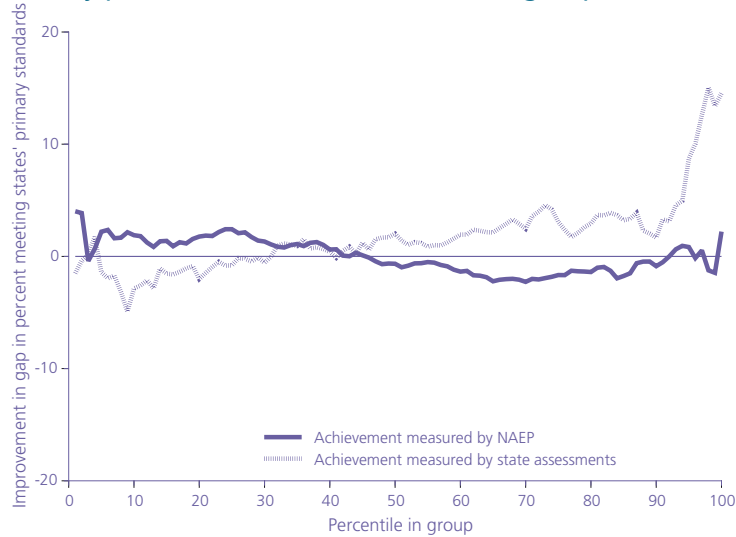


NOTE: Students eligible for free/reduced price lunch are referred to as (economically) disadvantaged. Percentile in group refers to the percentage of the disadvantaged (or non-disadvantaged) student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.



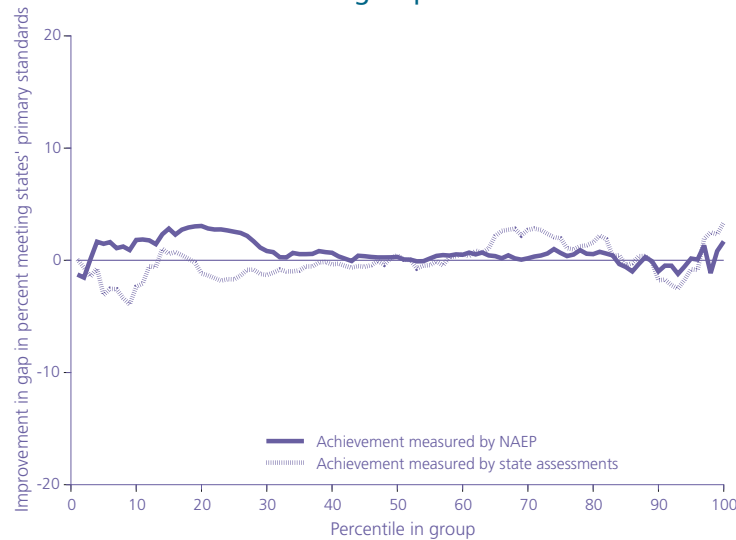
**Figure 34.** Profile of the reduction of the Hispanic-White gap in school percentage of students meeting states' primary grade 4 reading achievement standards, as measured by NAEP and state assessments, by percentile of students in each subgroup: 2002 and 2003



NOTE: Percentile in group refers to the percentage of the Hispanic or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

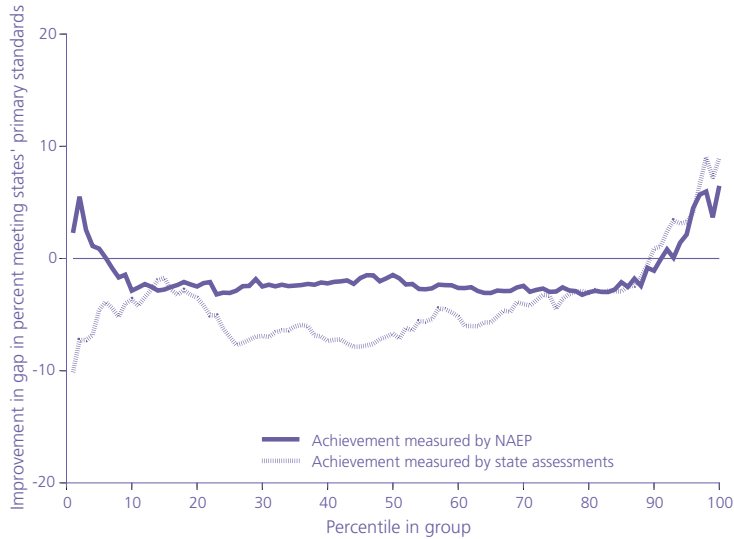
**Figure 35.** Profile of the reduction of the Black-White gap in school percentage of students meeting states' primary grade 4 reading achievement standards, as measured by NAEP and state assessments, by percentile of students in each subgroup: 2002 and 2003



NOTE: Percentile in group refers to the percentage of the Black or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.

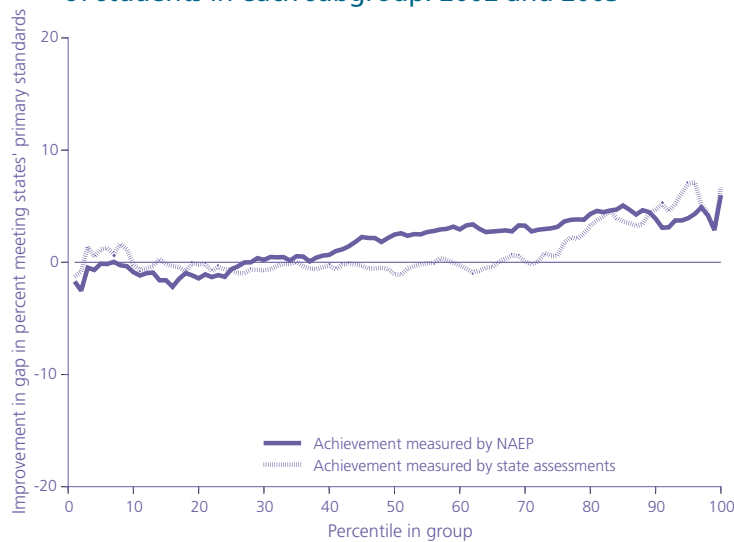
Figure 36. Profile of the reduction of the Hispanic-White gap in school percentage of students meeting states' primary grade 8 reading achievement standards, as measured by NAEP and state assessments, by percentile of students in each subgroup: 2002 and 2003



NOTE: Percentile in group refers to the percentage of Hispanic or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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 37. Profile of the reduction of the Black-White gap in school percentage of students meeting states' primary grade 8 reading achievement standards, as measured by NAEP and state assessments, by percentile of students in each subgroup: 2002 and 2003



NOTE: Percentile in group refers to the percentage of the Black or White student population who are in schools with lower (same-group) percentages meeting the states' primary reading standards.

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.





Profiles of grade 4 reading gap reduction between 2002 and 2003 for individual states are presented in appendix D. These profiles are accompanied by tables of mean discrepancies between NAEP and state assessment results, both for the overall populations compared and for segments of the subpopulations: the higher- and lower-achieving halves, the highest and lowest quarters, and the middle half of the achievement distributions. Student's *t* tests for the six comparisons are included, rather than simple indicators of statistical significance, to allow readers to judge significance based on their needs. Generally, for a single test in isolation, based on a large sample, a Student's *t* value greater than 1.96, either positive or negative, is considered statistically significant.

The differences between NAEP and state assessment measurement of the changes in state gaps in percentages meeting the primary grade 4 reading standard between 2002 and 2003 are shown in table 12 (for the states in which sufficient subgroup data are available). Of the 37 comparisons shown in table 12, only two Black-White gaps were significant at the .05 level: NAEP found significantly less Black-White gap reduction in 2003 than the state assessment did in Louisiana and Ohio. In addition, one finding related to poverty gaps was significant: NAEP found significantly more gap reduction in California between 2002 and 2003 than did the state assessment. At grade 8, as shown in table 13, the results were similar. In only one of the 32 comparisons was there a statistically significant difference: NAEP found significantly less poverty gap reduction in Hawaii.

Overall, the number of statistically significant differences in gap reductions is approximately the number expected by chance variation. This may be at least partially attributable to limitations in the database, which only contains one-year trends up to 2003. Furthermore, suppression of state assessment scores for schools with very small numbers of tested subgroup members reduces the sample size for these analyses (which are based on subsets of the schools selected for NAEP participation), with an accompanying reduction in the reliability of state mean gap estimates. Finally, trend results are subject to additional random error due to the selection of a different random sample of schools to participate in each NAEP assessment.

Table 12. Differences between NAEP and state assessments of grade 4 reading achievement gap reductions from 2002 to 2003, by state

State/ jurisdiction	Black-White	Hispanic-White	Poverty
Alabama	3.5	—	0.4
Alaska	—	—	—
Arizona	—	-3.4	—
Arkansas	-0.9	—	4.5
California	—	3.5	10.4 *
Colorado	—	—	—
Connecticut	-4.4	0.2	—
Delaware	3.0	—	-0.9
District of Columbia	—	—	—
Florida	—	—	—
Georgia	—	—	—
Hawaii	—	—	-5.0
Idaho	—	—	—
Illinois	3.3	-0.3	-0.5
Indiana	-8.1	—	-8.3
Iowa	—	—	—
Kansas	—	—	—
Kentucky	6.1	—	-2.8
Louisiana	-6.7 *	—	—
Maine	—	—	—
Maryland	—	—	—
Massachusetts	—	—	—
Michigan	—	—	—
Minnesota	—	—	—
Mississippi	-3.6	—	—
Missouri	6.1	—	—
Montana	—	—	—
Nebraska	—	—	—
Nevada	—	—	—
New Hampshire	—	—	—
New Jersey	—	—	—
New Mexico	—	—	—
New York	-1.4	-6.3	-0.8
North Carolina	-4.7	—	-3.5
North Dakota	—	—	—
Ohio	-14.4 *	—	—
Oklahoma	—	—	—
Oregon	—	—	—
Pennsylvania	-2.7	—	-4.7
Rhode Island	—	—	—
South Carolina	-2.7	—	—
South Dakota	—	—	—
Tennessee	0.9	—	3.1
Texas	6.1	3.2	—
Utah	—	—	—
Vermont	—	—	—
Virginia	—	—	—
Washington	—	—	—
West Virginia	—	—	1.2
Wisconsin	13.1	—	—
Wyoming	—	—	—

— Not available.

\* The NAEP-state difference is statistically significant at  $p < .05$ .

NOTE: A positive entry indicates that NAEP reports a larger gap reduction than state assessment does; a negative entry indicates NAEP reports a smaller gap reduction.

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.

Table 13. Differences between NAEP and state assessments of grade 8 reading achievement gap reductions from 2002 to 2003, by state

State/ jurisdiction	Black-White	Hispanic-White	Poverty
Alabama	4.3	—	0.7
Alaska	—	—	—
Arizona	—	-5.9	—
Arkansas	7.1	—	6.2
California	—	—	—
Colorado	—	—	—
Connecticut	1.3	4.1	—
Delaware	-0.9	—	-0.4
District of Columbia	—	—	—
Florida	—	—	—
Georgia	—	—	—
Hawaii	—	—	-6.4 *
Idaho	—	—	—
Illinois	-2.9	-1.7	-8.6
Indiana	2.7	—	-5.2
Iowa	—	—	—
Kansas	—	—	—
Kentucky	—	—	4.0
Louisiana	0.2	—	—
Maine	—	—	—
Maryland	—	—	—
Massachusetts	—	—	—
Michigan	—	—	—
Minnesota	—	—	—
Mississippi	7.6	—	—
Missouri	2.3	—	—
Montana	—	—	—
Nebraska	—	—	—
Nevada	—	—	—
New Hampshire	—	—	—
New Jersey	—	—	—
New Mexico	—	—	—
New York	-3.5	-6.5	-1.1
North Carolina	-1.9	—	-3.0
North Dakota	—	—	—
Ohio	—	—	—
Oklahoma	—	—	—
Oregon	—	—	—
Pennsylvania	-1.8	—	-5.5
Rhode Island	—	—	—
South Carolina	-3.7	—	—
South Dakota	—	—	—
Tennessee	-6.2	—	-0.0
Texas	7.1	1.3	—
Utah	—	—	—
Vermont	—	—	—
Virginia	—	—	—
Washington	—	—	—
West Virginia	—	—	-4.8
Wisconsin	—	—	—
Wyoming	—	—	—

— Not available.

\* The NAEP-state difference is statistically significant at  $p < .05$ .

NOTE: A positive entry indicates that NAEP reports a larger gap reduction than state assessment does; a negative entry indicates NAEP reports a smaller gap reduction.

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.

## SUMMARY

---

Comparisons are made between NAEP and state assessment measurements of (1) reading achievement gaps in grades 4 and 8 in 2003 and (2) changes in these reading achievement gaps between 2002 and 2003. Comparisons are based on school-level percentages of Black, Hispanic, White, and economically disadvantaged and non-disadvantaged students achieving the state's primary reading achievement standard in the NAEP schools in each state. In most states, the comparison is based on state test scores for grades 4 and 8, but where tests are not given at those grades, state assessment scores from adjacent grades are used for comparisons in a few states (table 4). Comparisons of gaps are subject to data availability. Generally, for subpopulations accounting for fewer than 10 percent of the student population in a state, reliable comparisons of gaps based on school-level data are not feasible due to suppression of small sample results.

Black-White gap comparisons for 2003 are possible in 26 states for grade 4 and 20 states for grade 8, Hispanic-White gap comparisons in 14 states for grade 4 and 13 states for grade 8, and poverty gap comparisons in 31 states for grade 4 and 28 states for grade 8. Gap reduction comparisons, which require scores for both 2002 and 2003, are possible for Black-White trends in 18 states for grade 4 and 15 states for grade 8, and poverty trends in 13 states for grade 4 and 12 states for grade 8. However, Hispanic-White trends can only be compared in 6 states for grade 4 and 5 states for grade 8.

In most states, gap profiles based on NAEP and state assessments are not significantly different from each other. Where differences were found, there is a tendency for NAEP to find slightly larger gaps than state assessments find. Of 132 state comparisons (composed of Black-White Hispanic-White, grade 4, and grade 8 comparisons across multiple states), NAEP found significantly larger gaps in 13 cases, while state assessments found significantly larger gaps in 8 cases. In those cases where NAEP found larger Black-White gaps, it was focused primarily in comparisons of the achievement of the lower-achieving half of the Black student population with the lower-achieving half of the White student population.

There was very little evidence of discrepancies between NAEP and state assessment measurement of gap reductions between 2002 and 2003. Aggregate gap reduction profiles based on NAEP and state assessments are very similar. Only four of 69 possible state comparisons were statistically significant, with NAEP indicating a smaller gap reduction in three of those cases.