

Achievement Gaps

How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress

Statistical Analysis Report



Achievement Gaps

How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress

Statistical Analysis Report

July 2009

Alan Vanneman
Linda Hamilton
Janet Baldwin Anderson
NAEP Education Statistics Services Institute

Taslina Rahman
National Center for Education Statistics

U.S. Department of Education

Arne Duncan
Secretary

Institute of Education Sciences

John Q. Easton
Director

National Center for Education Statistics

Stuart Kerachsky
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

July 2009

The NCES World Wide Web Home Page address is <http://nces.ed.gov>.
The NCES World Wide Web Electronic Catalog is <http://nces.ed.gov/pubsearch>.

Suggested Citation

Vanneman, A., Hamilton, L., Baldwin Anderson, J., and Rahman, T. (2009). *Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress*, (NCES 2009-455). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

For ordering information on this report, write to

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

or call toll free 1-877-4ED-Pubs or order online at <http://www.edpubs.org>.

Content Contact

Taslima Rahman
(202) 502-7316
taslima.rahman@ed.gov

The National Assessment of Educational Progress (NAEP), a congressionally mandated project of the U.S. Department of Education, informs the public periodically about the academic achievement of elementary and secondary students in reading, mathematics, science, writing and other subjects. Only information related to academic achievement and relevant variables is collected under this program from students representing the country. By making objective information available on performance of all race/ethnic groups at the national and state levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. While the National Center for Education Statistics (NCES) within the Institute of Education Sciences (IES) of the U.S. Department of Education conducts the survey, the National Assessment Governing Board oversees and sets policies for NAEP.

Executive Summary

In 2007, mathematics scores for both Black and White public school students in grades 4 and 8 nationwide, as measured by the main NAEP assessments of the National Assessment of Educational Progress (NAEP), were higher than in any previous assessment, going back to 1990. This was also true for Black and White fourth-graders on the NAEP 2007 Reading Assessment. For grade 8, reading scores for both Black and White students were higher in 2007 than in the first reading assessment year, 1992, as well as the most recent previous assessment year, 2005.

White students, however, had higher scores than Black students, on average, on all assessments. While the nationwide gaps in 2007 were narrower than in previous assessments at both grades 4 and 8 in mathematics and at grade 4 in reading, White students had average scores at least 26 points higher than Black students in each subject, on a 0-500 scale. This report will use results from both the main NAEP and the long-term trend NAEP assessments to examine the Black-White achievement gaps, and changes in those gaps, at the national and state level.

The main NAEP 2007 Reading and Mathematics Assessments included grade 4 and grade 8 students both nationally and for all 50 states, as well as the Department of Defense Education Activity (DoDEA) and the District of Columbia (hereinafter referred to as states). Not all states had Black (or White) student populations large enough to provide reliable data, and not all states participated in the earliest NAEP state assessments.

Most of the data in this report comes from the main NAEP assessments, supplemented with some data from the NAEP long-term trend assessments. Main NAEP assessments, which began in 1990 for mathematics and 1992 for reading, are administered at the fourth and eighth grades, both nationally and at the state level. Because main NAEP only

assesses public schools in its state assessments, this report contains only public school results. The most recent results in this report are for 2007.

NAEP long-term trend assessments are administered by age rather than grade. This report references long-term trend assessment public school results from the earliest assessment through 2004, with results for ages 9 and 13 instead of grades 4 and 8. The long-term trend assessments provide public school results for mathematics going back to 1978 and for reading going back to 1980, at ages 9, 13, and 17, at the national level only, on a 0-500 point scale.

At both ages 9 and 13, mathematics scores for both Black and White students were higher in 2004 than in any previous assessment. The 23-point Black-White achievement gap in mathematics for age 9 public school students in 2004 was narrower than in the first assessment in 1978 but not significantly different from the gap in the most recent previous assessment in 1999. The same was true for the 26-point gap at age 13.

For age 9 reading, scores for both Black and White students were higher in 2004 than in any previous assessment, going back to 1980. The 26-point gap between Black and White students in 2004 was not significantly different from the gap in 1980, but was narrower than the gap in 1999. At age 13 reading, scores were higher for Black students in 2004 than in 1980, but did not show a significant difference from 1999. Scores for White students were not significantly different for either comparison year. The 21-point gap in student performance at age 13 reading in 2004 was narrower than in both 1980 and 1999.

The following two sections summarize state-level achievement gaps between Black and White students in the main NAEP assessments in mathematics and reading.

State Black-White Achievement Gaps—Mathematics

- At the state level, gaps in grade 4 mathematics existed in 2007 in the 46 states for which results were available. In 15 states, the 2007 gaps were narrower than in 1992, as Black students demonstrated a greater gain in average scores than that of the White students.
- At grade 8, mathematics gaps existed in 2007 in the 41 states for which results were available. The gaps were narrower in 2007 than in 1990 in four states: Arkansas, Colorado, Oklahoma, and Texas. In all four, scores for both Black and White students increased, but scores for Black students increased more.
- At grade 4, five states had mathematics gaps in 2007 that were larger than the national gap of 26 points, while 10 states had gaps that were smaller.
- At grade 8, seven states had mathematics gaps in 2007 that were larger than the national gap of 31 points, while 12 had gaps that were smaller.

State Black-White Achievement Gaps—Reading

- At the state level, gaps in grade 4 reading existed in 2007 in the 44 states for which results were available. Gaps narrowed from 1992 to 2007 in Delaware, Florida, and New Jersey, due to larger increases in Black students' scores.
- At grade 8, reading gaps existed in 2007 in 41 of the 42 states for which results were available. In Hawaii, the 7-point difference between Black and White students' scores in 2007 was not statistically significant, and thus there was no gap for Hawaii. There was no significant change in the gap in any state from 1998 to 2007.

- At grade 4, eight states had reading gaps that were larger than the 2007 national gap of 27 points, while nine had gaps that were smaller.

- At grade 8, one state had a reading gap that was larger than the 2007 national gap of 26 points, while nine had gaps that were smaller.

The NAEP reading and mathematics scales make it possible to examine relationships between students' performance and various background factors measured by NAEP, such as race. However, a relationship that exists between achievement and another variable does not reveal its underlying cause, which may be influenced by a number of other variables. Similarly, the assessments do not reflect the influence of unmeasured variables. At the state level, changes in the size of the achievement gap between Black and White students could be affected by demographic changes in the size and makeup of the populations involved, as well as policy changes in the schools and communities. The results of this study are most useful when they are considered in combination with other knowledge about the student population and the education system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

This report focuses on the size of the achievement gap between Black and White students and the direction of average scores within states, regardless of the states' scores. Large gaps may occur in some states with scores above the national average, as well as in states with scores below the national average. Similarly, small gaps may occur in states with scores above or below the national average. All differences discussed in this report are statistically significant at the .05 level after controlling for multiple comparisons. The technical notes for this report provide information about sampling, accommodations, interpreting statistical significance, and other technical features. For more information on both the main NAEP and long-term trend assessments, see appendix A.

Contents

Executive Summary	iii
List of Tables	vii
List of Figures	viii
Introduction	1
Sources of the Main NAEP data	2
Sources of the Long-Term Trend NAEP data	3
Understanding score gaps	4
Understanding statistical significance	4
Cautions in interpreting the data	5
How this report is organized	5
MATHEMATICS	
Long-Term Trend Results for Black and White 9- and 13-Year-Olds	6
Trends in mathematics scores and achievement gaps, 1978–2004	6
Main NAEP National Results for Black and White Fourth- and Eighth-Graders	7
Trends in mathematics scores and achievement gaps, 1990–2007	7
Mathematics scores and achievement gaps by gender, 1990–2007	8
Mathematics scores and achievement gaps by family income, 2003–2007	10
Main NAEP State Results for Black and White Fourth- and Eighth-Graders	12
State and national mathematics achievement gaps at grade 4, 2007	13
Trends in state mathematics achievement gaps at grade 4, 1992–2007	14
State and national mathematics achievement gaps at grade 8, 2007	21
Trends in state mathematics achievement gaps at grade 8, 1990–2007	22
READING	
Long-Term Trend Results for Black and White 9- and 13-Year-Olds	28
Trends in reading scores and achievement gaps, 1980–2004	28
Main NAEP National Results for Black and White Fourth- and Eighth-Graders	29
Trends in reading scores and achievement gaps, 1992–2007	29
Reading scores and achievement gaps by gender, 1992–2007	30
Reading scores and achievement gaps by family income, 2003–2007	32
Main NAEP State Results for Black and White Fourth- and Eighth-Graders	34
State and national reading achievement gaps at grade 4, 2007	35
Trends in state reading achievement gaps at grade 4, 1992–2007	36
State and national reading achievement gaps at grade 8, 2007	43
Trends in state reading achievement gaps at grade 8, 1998–2007	44
Appendix A: Technical Notes	50
Frameworks, development, administration, scoring, and analysis	50
Sources of the data	50
NAEP sampling procedures	50
Understanding NAEP reporting groups	55
Race/ethnicity	55
Eligibility for free/reduced-price school lunch	56

Gender	56
Inclusion and exclusion	56
Accommodations	60
Drawing inferences from the results	60
Weighting and variance estimation	61
Analyzing group differences in averages and percentages	61
Conducting multiple tests	63
Cautions in interpretation	63
Appendix B: Supplemental Tables	64
Acknowledgments	66

List of Tables

Table	Page
1. Percentage of public school students assessed in NAEP mathematics by eligibility for free or reduced-price school lunch, race/ethnicity and grade: 2003, 2005, and 2007	10
2. Percentage of public school students assessed in NAEP reading by eligibility for free or reduced-price school lunch, race/ethnicity and grade: 2003, 2005, and 2007	32
A-1. School and student participation rates, and target student population, Long-Term Trend Reading assessment, public school students only, by age: 2004	51
A-2. School and student participation rates, and target student population, grade 8 reading assessment, public school students only, by state or jurisdiction: 2007	54
A-3. National Long-Term Trend mathematics and reading exclusion rates as percentages of the total sample, by age, type of assessment and race/ethnicity: 2004	57
A-4. National mathematics and reading exclusion rates as percentages of the total sample, public schools only, by grade and race/ethnicity: 2007	58
A-5. Mathematics and reading exclusion rates as percentages of the total sample, public schools only, by grade, race/ethnicity and jurisdiction: 2007	59
B-1. Administration of NAEP national and state mathematics assessments, by grade: Various years, 1990–2007	64
B-2. Average mathematics scale scores for all public school students at grades 4 and 8, by gender and eligibility for the National School Lunch Program: Various years, 1990–2007	64
B-3. Administration of NAEP national and state reading assessments, by grade: Various years, 1992–2007	65
B-4. Average reading scale scores for all public school students at grades 4 and 8, by gender and eligibility for the National School Lunch Program: Various years, 1992–2007	65

List of Figures

Figure	Page
1. Trends in average mathematics scale scores and score gaps for White students and Black students at age 9: Various years, 1978–2004	6
2. Trends in average mathematics scale scores and score gaps for White students and Black students at age 13: Various years, 1978–2004	6
3. Mathematics achievement score gaps between Black and White public school students at grade 4: Various years, 1990–2007	7
4. Mathematics achievement score gaps between Black and White public school students at grade 8: Various years, 1990–2007	7
5. Gaps in average mathematics scores between Black and White public school students at grade 4, by gender: Various years, 1990–2007	9
6. Gaps in average mathematics scores between Black and White public school students at grade 8, by gender: Various years, 1990–2007	9
7. Gaps in average mathematics scores between Black and White public school students at grade 4, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007	11
8. Gaps in average mathematics scores between Black and White public school students at grade 8, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007	11
9. The Black-White achievement score gap in mathematics for public school students at grade 4, by state or jurisdiction: 2007	13
10. Gaps in average mathematics scores between Black and White public school students at grade 4, by state: Various years, 1992–2007	15
11. The Black-White achievement score gap in mathematics for public school students at grade 8, by state or jurisdiction: 2007	21
12. Gaps in average mathematics scores between Black and White public school students at grade 8, by state: Various years, 1990–2007	23
13. Trends in average reading scale scores and score gaps for White students and Black students at age 9: Various years, 1980–2004	28
14. Trends in average reading scale scores and score gaps for White students and Black students at age 13: Various years, 1980–2004	28
15. Reading achievement score gaps between Black and White public school students at grade 4: Various years, 1992–2007	29
16. Reading achievement score gaps between Black and White public school students at grade 8: Various years, 1992–2007	29
17. Gaps in average reading scores between Black and White public school students at grade 4, by gender: Various years, 1992–2007	31
18. Gaps in average reading scores between Black and White public school students at grade 8, by gender: Various years, 1992–2007	31
19. Gaps in average reading scores between Black and White public school students at grade 4, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007	33

Figure	Page
20. Gaps in average reading scores between Black and White public school students at grade 8, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007	33
21. The Black-White achievement score gap in reading for public school students at grade 4, by state or jurisdiction: 2007	35
22. Gaps in average reading scores between Black and White public school students at grade 4, by state: Various years, 1992–2007.	37
23. The Black-White achievement score gap in reading for public school students at grade 8, by state or jurisdiction: 2007	43
24. Gaps in average reading scores between Black and White public school students at grade 8, by state: Various years, 1998–2007.	45



Introduction

The past half century has witnessed considerable gains in educational attainment in the United States. Between 1950 and 2005, the percentage of young adults ages 25-29 who had completed high school rose from 53 to 86. For White young adults, the percentage increased from 56 to 93, and for Black young adults it increased from 24 to 86.¹

There have also been gains in educational achievement. National and state mathematics scores in grades 4 and 8 on the National Assessment of Educational Progress (NAEP) were at their highest levels in 2007.² Reading scores for the nation and a substantial number of states have also increased since the early 1990s.³

Although scores have increased for both Black students and White students, on average Black students do not perform as well as their White peers. At the national level, the fourth-grade Black-White achievement gap in mathematics for 2007 was narrower than in 1990, while the fourth-grade reading gap was narrower than in either 1992 or 2005. At the eighth grade, the gap in mathematics was narrower in 2007 than in 2005, while the reading gap did not change significantly compared to either prior assessment year.

The Elementary and Secondary Education Act⁴ when first authorized intended to improve the educational achievement of low-performing students, particularly low-income students and Black students. Subsequent reauthorizations of the act have reaffirmed the importance of closing the achievement gaps. This report uses NAEP data to examine the progress of the nation and each of the states in reducing the gap between Black and White students at grades 4 and 8 in both reading and mathematics.

Issues relating to the Black-White achievement gap have been addressed by a number of recent studies. *Status and Trends in the Education of Racial and Ethnic Minorities*,⁵ issued by the National Center for Education Statistics (NCES), for example, examined the education of all major racial and ethnic groups in the United States from pre-kindergarten through the postsecondary level, along with employment and income data for these groups. The report identified a variety of factors which are correlated with the achievement gap between Black and White students. For example, Black students were more likely than White students to come from families living in poverty, which is associated with lower educational performance.

Other reports have used NAEP data in analyses attempting to isolate important factors related to the Black-White achievement gap. For example, *The Family: America's Smallest School*,⁶ issued by the Educational Testing Service, correlates student achievement, as measured by NAEP, with four home factors: the presence of two parents in the home, the hours children spend watching television, the hours parents spend reading to them, and the frequency of absence from school. Compared to White students, Black children were less likely to come from a family with both parents in the home, spent more hours watching television, were read to by their parents for fewer hours, and were more likely to be absent from school.

Another report issued by the Educational Testing Service, *Parsing the Achievement Gap II*,⁷ considered 16 factors previously identified as being correlated with how well students performed in school. Seven were school-related (including, for example, curriculum rigor and teacher preparation), eight “before and after” school factors (including, for example, weight at birth, exposure to lead, and excessive TV watching), and the “home school connection,” parent

¹ Snyder, T.D., Dillow, S.A., and Hoffman, C.M. (2007). *Digest of Education Statistics 2006* (NCES 2007-017). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

² Lee, J., Grigg, W., and Dion, G. (2007). *The Nation's Report Card: Mathematics 2007* (NCES 2007-494). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, D.C.

³ Lee, J., Grigg, W., and Donahue, P. (2007). *The Nation's Report Card: Reading 2007* (NCES 2007-496). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, D.C.

⁴ Elementary and Secondary Education Act of 1965, P.L. 89-10, 79 Stat. 27.

⁵ KewalRamani, A., Gilbertson, L., Fox, M., and Provasnik, S. (2007). *Status and Trends in the Education of Racial and Ethnic Minorities* (NCES 2007-039). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

⁶ Barton, P., and Coley, R. (2007). *The Family: America's Smallest School*. Princeton, NJ: Educational Testing Service.

⁷ Barton, P., and Coley, R. (2009). *Parsing the Achievement Gap II*. Princeton, NJ: Educational Testing Service.

participation. Using data from NAEP and other sources, the report said that for all 16 factors there were gaps that favored White students over Black students—for example, White students were more likely than Black students to attend schools offering rigorous curriculums and less likely to suffer from low birth weight.

This report uses data from both the “main NAEP” and the NAEP long-term trend assessments. NCES and the National Assessment Governing Board, which sets policy for NAEP, have maintained comparability of data for both main and long-term trend NAEP. Main NAEP assessments, which began in 1990 for mathematics and 1992 for reading, are administered at the fourth and eighth grades, both nationally and at the state level. The most recent administration was in 2007. The long-term trend assessments provide public school results for mathematics going back to 1978 and for reading going back to 1980, at ages 9, 13, and 17, at the national level only. The most recent long-term trend report available at the time of the preparation of this report contains results for the assessments administered in 2004. Discussion of main NAEP grade 12 assessments is omitted in this report because these assessments are conducted at the national level only.

While the main NAEP assessments do not go as far back in time as the long-term trend assessments, they allow the examination of trends in the Black and White performance gap in every state, plus the District of Columbia and the U.S. Department of Defense Education Activity (DoDEA) schools. In addition, the main NAEP assessments use frameworks that are more closely aligned with current practices regarding instructional content; they include more questions overall and more questions that require a written response; and they employ much larger samples than long-term NAEP.

All data presented in this report for main NAEP are for public school students only. Main NAEP and long-term trend provide national results for both public and private school students, but NAEP state results are for public school students only. To maintain consistency of data for comparison purposes, this report uses only public school data at the national level as well.

The major questions addressed in this study are: 1) how do gaps in 2007 compare to the gaps in the initial and most recent prior years of the NAEP national and state assessment series? And 2) how do states compare to the nation in 2007? The current report presents these results in graphs that show the NAEP achievement gaps in a format that makes it possible to see at a glance the national and state gaps results for all available years.

In previous NAEP reports, achievement gaps results have been available to users in two ways: 1) online, using the NAEP Data Explorer, and 2) by year, in the report cards for a given assessment. The NAEP Achievement Gaps report is the first NCES publication to present the Black and White NAEP achievement gaps across time for all the states and the nation, including results for every assessment year since state assessments began.

States first participated in the eighth-grade mathematics assessment in 1990, the fourth-grade reading and mathematics assessments in 1992, and the eighth-grade reading assessment in 1998. The No Child Left Behind Act of 2001 requires each state, beginning in 2003, to participate in the NAEP mathematics and reading assessments if they are to receive Title I education funding (Public Law 107-110 Title I Part A, Sec. 1111). Prior to the passage of the Act, participation was voluntary and about 40 states participated in each assessment. (In this report, “state” and “jurisdiction” will be used interchangeably to refer to the 50 states, the District of Columbia, and the Department of Defense Education Activity (DoDEA) schools.) Additional information about the years when the national and state assessments were administered is in appendix B.

Sources of the Main NAEP data

This report presents national data from the NAEP reading and mathematics assessments for Black and White public school students at the fourth and eighth grades. Only results for White (non-Hispanic) and Black (non-Hispanic) public school students are contained in this report. Additional information on the national and state assessments is given in appendix B.

Administration of main NAEP national and state reading and mathematics assessments

			1990	1992	1994	1996	1998	2000	2002	2003	2005	2007
Reading	4th Grade	National		✓	✓		✓	✓	✓	✓	✓	✓
		State		✓	✓		✓		✓	✓	✓	✓
	8th Grade	National		✓	✓		✓	✓	✓	✓	✓	✓
		State					✓		✓	✓	✓	✓
Mathematics	4th Grade	National	✓	✓		✓		✓		✓	✓	✓
		State		✓		✓		✓		✓	✓	✓
	8th Grade	National	✓	✓		✓		✓		✓	✓	✓
		State	✓	✓		✓		✓		✓	✓	✓

In 2007, Black and White students together comprised about three-fourths of the nation’s public school students at the fourth and eighth grades. At the fourth-grade level, 58 percent of assessed students were White and 16 percent were Black. At the eighth-grade level, 60 percent of assessed students were White and 16 percent were Black.

In the earliest main NAEP assessments, students with disabilities and English language learners did not receive accommodations. Since 1994 (1996 at the state level) students receiving accommodations on their state assessment received the same accommodations on NAEP, as long as NAEP approved them (see appendix A for details.)

In 2007, the reading assessment was given to 183,000 fourth-graders and 155,000 eighth-graders, while the mathematics assessment was given to 190,000 fourth-graders and 147,000 eighth-graders. The main NAEP samples are so large because they include representative samples for each of the 50 states, plus the District of Columbia and Department of Defense school system for Armed Forces dependents in the United States and overseas. This allows examination of the achievement gaps for public school students for individual states as well as for the nation as a whole.

NAEP assessments are conducted in a six-week window starting in January of each assessment year. The same assessment is administered in both the national and state assessments. Because the content of the assessments given to fourth-graders and eighth-graders differs, scores for the two grades should be compared with caution, even though the scores appear on similar 0-500 scales. Scores for reading and mathematics cannot be compared because the two assessments are scaled independently. See appendix A for more details.

Sources of the Long-Term Trend NAEP data

This report presents national data for public school students aged 9 and 13 from the 1978, 1982, 1986, 1990, 1992, 1994, 1996, 1999, and 2004 mathematics long-term trend assessments and the 1982, 1986, 1990, 1992, 1994, 1996, 1999, and 2004 reading long-term trend assessments. Unlike the main NAEP assessments, these assessments did not allow accommodations for students with disabilities and English language learners for the years included in this report. Sample sizes for the 2004 long-term trend assessments were 7,500 (9-year-old students) and 8,300 (13-year-old students) for reading and 7,300 (9-year-old students) and 7,500 (13-year-old students) for mathematics. See appendix A for more details.

Understanding score gaps

Ways the gap can change

The achievement gap between Black and White students is defined as the difference between the average score for Black students and the average score for White students. Comparisons are made for main NAEP between the most recent assessment year (2007) and all previous assessment years. Only changes between the earliest assessment year and 2007, and between 2005 and 2007, are discussed. For long-term trend, only changes between the earliest assessment year and 2004, and between 1999 and 2004, are discussed.

Changes in the size of the achievement gap depend on both changes in the average scores for Black and White students and the rate of change in those scores. Generally, widening gaps are seen as undesirable, while narrowing gaps are seen as desirable. However, it is possible for the gap to widen even if scores for both Black students and White students increase, if scores for the higher scoring group increase more than scores for the other group. And it is also possible for the gap to narrow even if scores for both Black and White students decline, if scores for the higher scoring group decline more than those of the other group. The following images illustrate the various ways that gaps can narrow.

It is important to note that although NAEP data can identify gaps and changes in gaps, these data cannot explain why gaps exist or why they change. NAEP assessments are designed to measure student performance and identify factors associated with it, not to identify or explain the causes of differences in student performance.

Understanding statistical significance

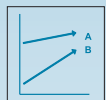
NAEP data are based on samples of students, and the results are subject to sampling and measurement error. Statistical tests are used to determine whether the differences between average scores are statistically significant—that is, whether they exceed the margin of error. Changes in average scores for Black students and White students and changes in the size of the gap between these scores are analyzed separately. Therefore, it is possible for the size of the achievement gap to increase or decrease even though the average scores of neither Black nor White students changed statistically significantly during the same period.

The term “significant” is not intended to imply a judgment about the absolute magnitude or the educational relevance of the differences. It is intended to identify statistically reliable population differences to help inform discussion among policymakers, educators, researchers, and the public.

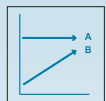
Beginning in 2002, the main NAEP national sample was obtained by aggregating the samples from each state, rather than by using an independently selected national sample. As a result, the national samples in mathematics and reading were larger in 2003, 2005, and 2007 than in previous assessment years. Thus, smaller score differences between years or between student groups were found to be statistically significant than would have been detected in previous assessments. All differences discussed in the text are significant at the .05 level with appropriate adjustments for part-to-whole and multiple comparisons.

Statistical comparisons of NAEP scores from different assessment years are made using a multiple comparison procedure (see appendix A, “Conducting multiple tests,” for details). However, in figures 9, 11, 21, and 23, comparisons of the size of the Black-White achievement gap

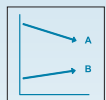
Ways gaps can narrow



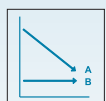
The average scores of *both* groups increase, while the score of the lower performing group increases even more.



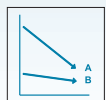
The average score of the *higher* performing group does not change, while the score of the lower performing group increases.



The average score of the *higher* performing group declines, while the score of the lower performing group increases.



The average score of the *higher* performing group declines, while the score of the lower performing group does not change.



The average scores of *both* groups decline, but the score of the higher performing group declines even more.

for each state to the national gap are made using pairwise comparisons, where each state is compared to the nation one at a time. For this reason, the results shown in these four figures may not correspond to results obtained from the NAEP Online Data Tool, which currently does not permit pairwise comparisons for this type of gap analysis.

Cautions in interpreting the data

All results given here are in terms of average scores, which reflect a wide range of student performance. Many Black students score above the average for White students and many White students score below the average for Black students. For detailed information on variations in performance, including standard deviations, consult the NAEP Data Explorer online at <http://nces.ed.gov/nationsreportcard/nde/viewresults.asp>

The analysis of NAEP data contained in this report should not be seen to imply causal relations. Simple cross-tabulations of a variable with measures of educational achievement, like the ones presented here, cannot be considered as evidence that differences in the variable cause differences in educational achievement. As noted earlier, NAEP surveys are not designed to identify causal relationships. There are many possible reasons why the performance of one group of students

will differ from that of another. Inferences related to student group performance should take into consideration the many socioeconomic and educational factors that may also be associated with performance.

All statistical tests are performed using unrounded scale scores. The Black-White achievement gap is calculated by subtracting the average scale score for Black students from the average scale score for White students. Because all results are presented as rounded numbers, occasionally the lower scale score plus the gap will not equal the higher scale score shown in this report's graphics.

How this report is organized

The remainder of this report presents first mathematics and then reading results. In each section, long-term trend results are presented first, giving national results only for public school students ages 9 and 13. These are followed by both national and state results for public school fourth- and eighth-graders from main NAEP. National data from main NAEP are also presented by 1) gender and 2) eligibility categories for the National School Lunch Program. The last section consists of an appendix that contains relevant technical notes and supplemental tables.

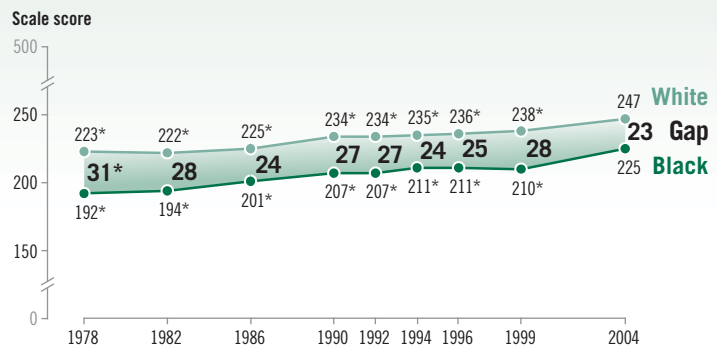


Long-Term Trend Results for Black and White 9- and 13-Year-Olds

Trends in mathematics scores and achievement gaps, 1978–2004

Mathematics scores for both 9- and 13-year-old Black and White students were higher in 2004 than on any previous long-term trend assessment (figures 1 & 2). In addition, the score gaps for Black and White students were narrower in 2004 than in the first assessment in 1978 for both age groups, as scores of Black students showed a greater increase than those of White students. The gaps in 2004 were not significantly different from the gaps in 1999.

Figure 1. Trends in average mathematics scale scores and score gaps for White students and Black students at age 9: Various years, 1978–2004

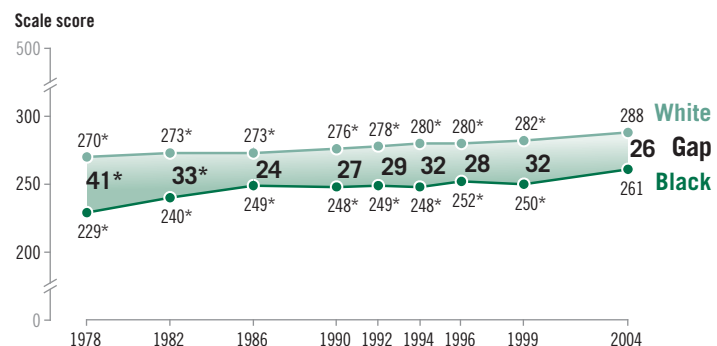


* Significantly different ($p < .05$) from 2004.

NOTE: Detail may not sum to totals due to rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1978–2004 Long-Term Trend Mathematics Assessments.

Figure 2. Trends in average mathematics scale scores and score gaps for White students and Black students at age 13: Various years, 1978–2004



* Significantly different ($p < .05$) from 2004.

NOTE: Detail may not sum to totals due to rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1978–2004 Long-Term Trend Mathematics Assessments.

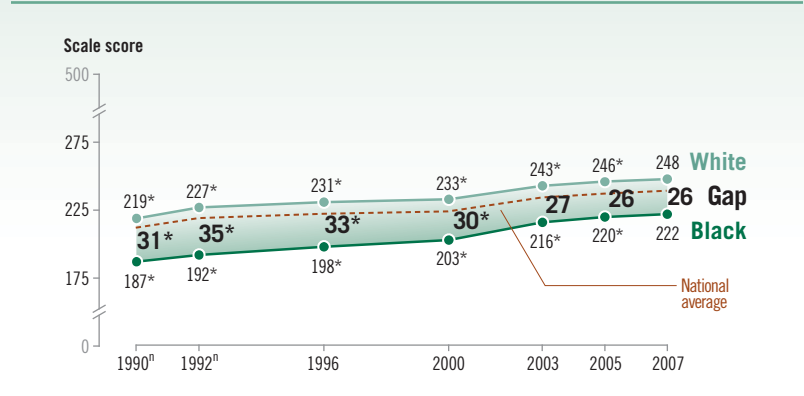
Main NAEP National Results for Black and White Fourth- and Eighth-Graders

Trends in mathematics scores and achievement gaps, 1990–2007

In main NAEP, average fourth-grade mathematics scores for the nation were higher in 2007 than in 1990 for both Black and White public school students (figure 3). The greater increase for Black fourth-graders resulted in the gap narrowing from 31 points in 1990 to 26 points in 2007. From 2005 to 2007, scores increased for both Black and White students, but there was no significant change in the gap.

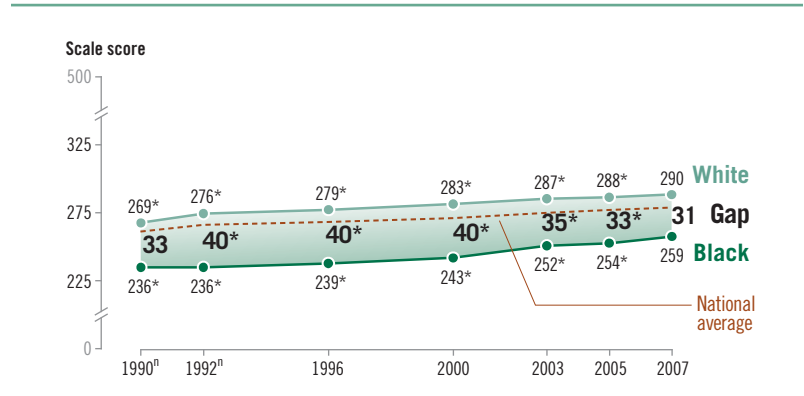
Average mathematics scores were higher in 2007 than in 1990 for both Black and White eighth-graders (figure 4). The 31-point gap in 2007 was not significantly different from the 33-point gap in 1990. However, the gap was narrower in 2007, at 31 points, than in 2005, at 33 points. Although scores for both groups were higher in 2007, a greater increase in Black students' scores caused the gap to narrow. The 2-point decrease in the gap from 2005 to 2007 was significant while the 2-point decrease from 1990 to 2007 was not. It is possible that the smaller standard errors in 2005, due to the increased sample size in that year, allowed the difference in 2005 to be identified as statistically significant.

Figure 3. Mathematics achievement score gaps between Black and White public school students at grade 4: Various years, 1990–2007



ⁿ Accommodations were not permitted for this assessment.
 * Significantly different (p<.05) from 2007.
 NOTE: Detail may not sum to totals due to rounding.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1990–2007 Mathematics Assessments.

Figure 4. Mathematics achievement score gaps between Black and White public school students at grade 8: Various years, 1990–2007



ⁿ Accommodations were not permitted for this assessment.
 * Significantly different (p<.05) from 2007.
 NOTE: Detail may not sum to totals due to rounding.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1990–2007 Mathematics Assessments.

Mathematics scores and achievement gaps by gender, 1990–2007

Average mathematics scores were higher in 2007 than in 1990 for the nation's Black and White fourth-graders, regardless of gender (figure 5). Among females, the gap was narrower in 2007 as the average score gains of Black females were greater than those of their White peers. Among fourth-grade males, the Black-White gap did not change significantly.

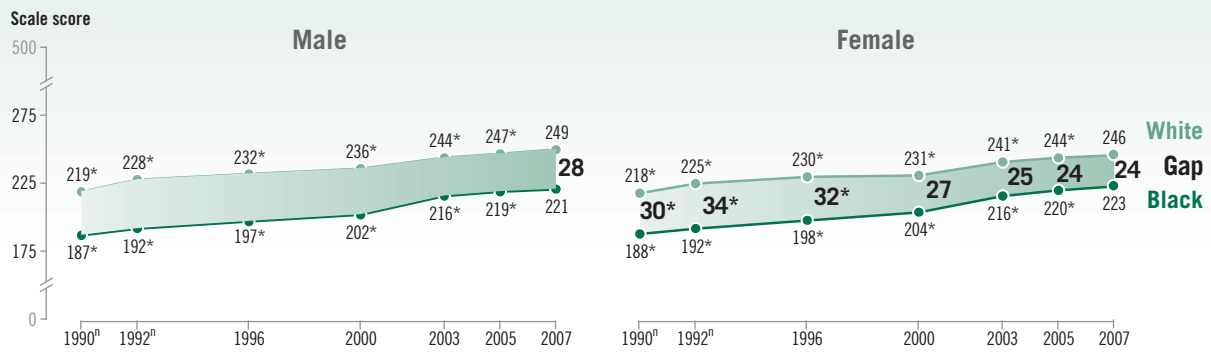
In addition to the 17-year gain, mathematics scores also increased during the two-year period, 2005 to 2007, for both Black and White fourth-graders, regardless of gender. However, the gaps did not change significantly either for males or for females during this period.

In 2007, average mathematics scores were higher than they had been in 1990 for Black and White eighth-graders (figure 6). However, the Black-White mathematics gap did not change significantly for either males or females.

At grade 8, mathematics scores increased from 2005 to 2007 for Black and White students, regardless of gender (figure 6). Female eighth-graders showed a narrowing of the gap during this period as Black females' scores increased more than those of White females, while the gap for males did not change significantly.



Figure 5. Gaps in average mathematics scores between Black and White public school students at grade 4, by gender: Various years, 1990–2007

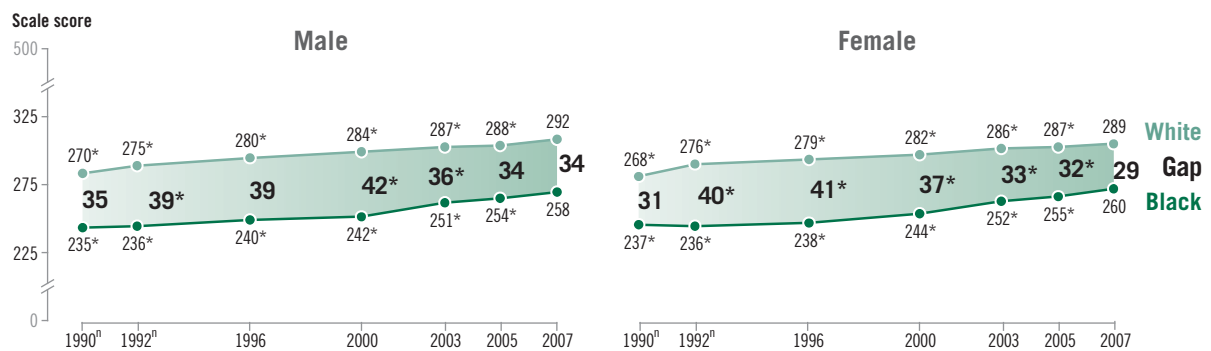


ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1990–2007 Mathematics Assessments.

Figure 6. Gaps in average mathematics scores between Black and White public school students at grade 8, by gender: Various years, 1990–2007



ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1990–2007 Mathematics Assessments.

Mathematics scores and achievement gaps by family income, 2003–2007

NAEP uses student eligibility for free or reduced-price school lunch as an indicator of family income. At grade 4, mathematics scores were higher in 2007 than in 2003 and 2005 for all Black and White public school students, regardless of school-lunch eligibility (figure 7). Despite these increases, the only significant Black-White gap change was between 2003 and 2007, for students eligible for reduced-price lunch.

At grade 8, mathematics scores were higher in 2007 than in 2003 and 2005 for all Black and White public school students (figure 8). The Black-White score gaps for students eligible for free or reduced-price lunch narrowed in 2007 in comparison to both previous assessments, as scores for eligible Black students showed greater gains than those of their White peers.

Eligibility for free and reduced-price lunch

NAEP collects data on students' eligibility for the National School Lunch Program (NSLP)—sometimes referred to as the free and reduced-price school lunch program—as an indicator of family economic status. Eligibility for free and reduced-price lunch is based on students' family income in relation to the federally established poverty level.

Not eligible: Students who are not eligible for the program because their family's income is above 185 percent of the poverty level.

Eligible for reduced-price lunch: Students who are eligible for reduced-price lunch because their family's income is between 130 percent and 185 percent of the poverty level.

Eligible for free lunch: Students who are eligible for free lunch because their family's income is below 130 percent of the poverty level.

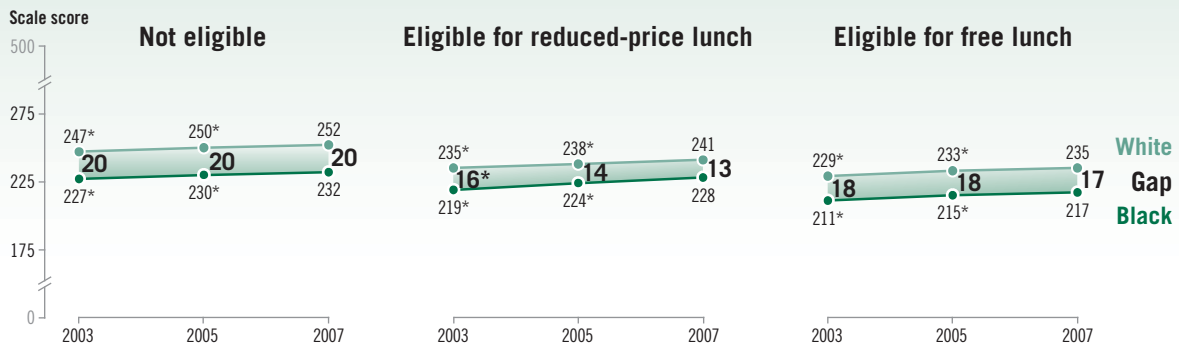
As a result of improvements in the quality of the data on students' eligibility for NSLP, the percentage of students for whom information was not available has decreased in comparison to the percentages reported prior to the 2003 assessment. Therefore, trend comparisons are only made back to 2003 in this report.

Table 1. Percentage of public school students assessed in NAEP mathematics by eligibility for free or reduced-price school lunch, race/ethnicity and grade: 2003, 2005, and 2007

	Not eligible		Eligible for reduced-price lunch		Eligible for free lunch	
	Black	White	Black	White	Black	White
Grade 4						
2007	26	72	7	6	66	21
2005	25	71	8	7	66	20
2003	24	72	9	8	66	19
Grade 8						
2007	32	76	7	5	60	18
2005	31	75	9	6	58	17
2003	32	76	9	6	56	15

NOTE: Detail may not sum to totals due to rounding.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003, 2005, and 2007 Mathematics Assessments.

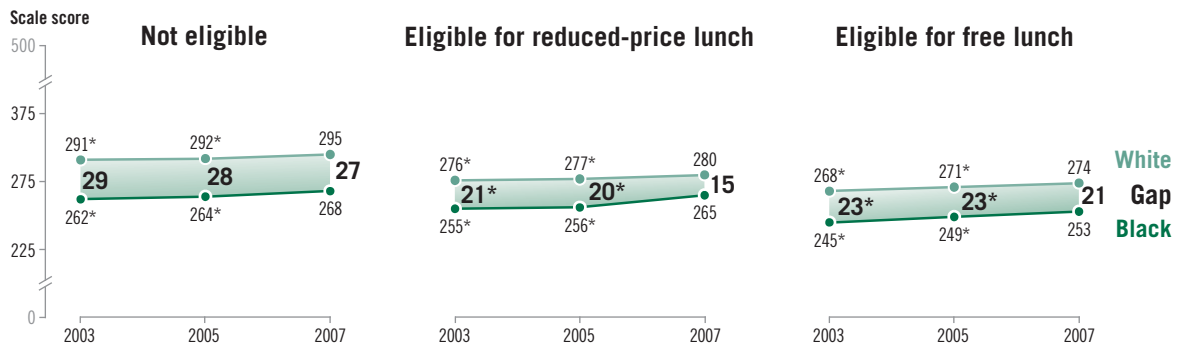
Figure 7. Gaps in average mathematics scores between Black and White public school students at grade 4, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007



* Significantly different (p<.05) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003, 2005, and 2007 Mathematics Assessments.

Figure 8. Gaps in average mathematics scores between Black and White public school students at grade 8, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007



* Significantly different (p<.05) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003, 2005, and 2007 Mathematics Assessments.

Main NAEP State Results for Black and White Fourth- and Eighth-Graders

The NAEP state mathematics assessments were administered to public school fourth-graders in 1992, 1996, 2000, 2003, 2005, and 2007 and to public school eighth-graders in 1990, 1992, 1996, 2000, 2003, 2005, and 2007. Before 2003, states were not required to participate in NAEP to qualify for Title I education funds. Typically, 40 or more states participated in each prior assessment. In 2003, 2005, and 2007, all 50 states, the District of Columbia, and the DoDEA participated.

State results are presented in two ways. Comparisons of fourth-grade mathematics gaps in 2007 between each state and the nation are presented in figure 9.

Comparisons of the mathematics gaps within a state over time are presented in a series of small graphs in figure 10. At the top left of each two-page spread, the mathematics scores and gaps for the nation are presented for reference. Each state figure, as well as the national figure, also contains a dotted red line representing the national average for public school students. The data for the national averages are located in the appendix in Table B-2.

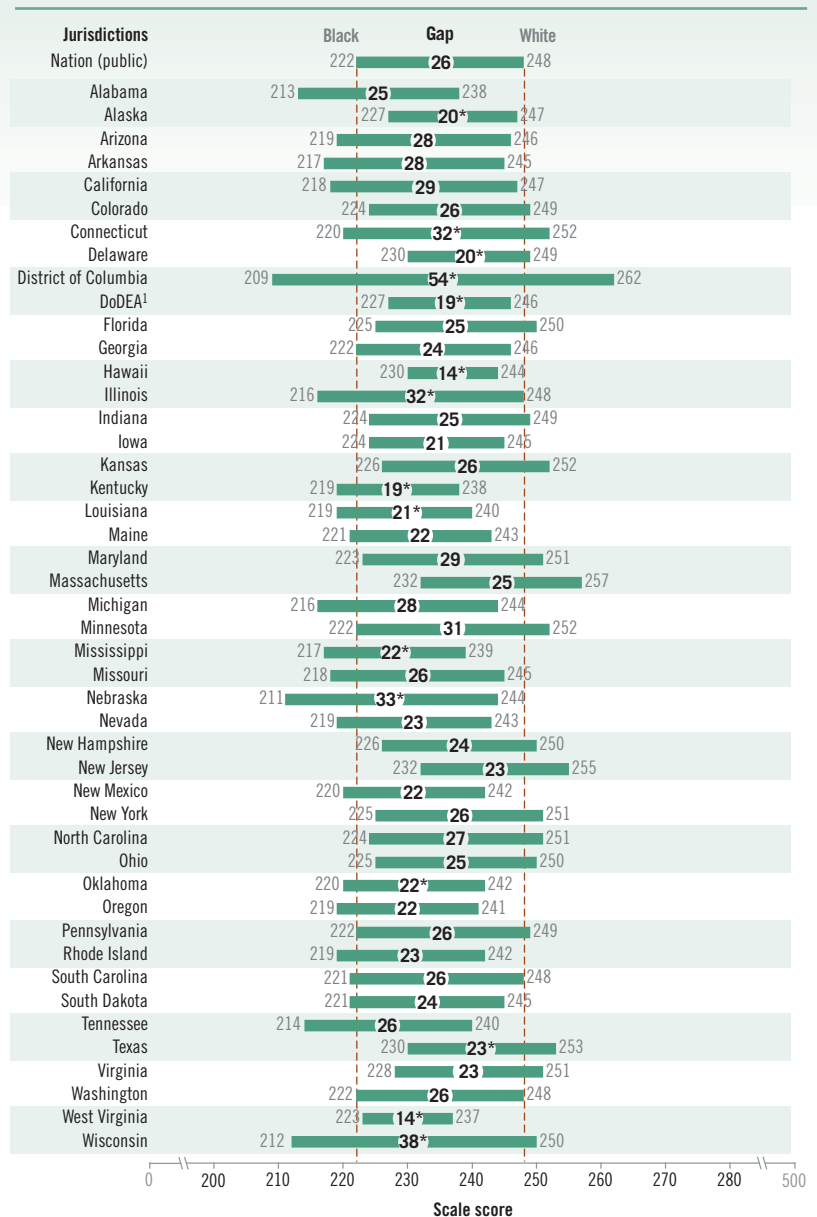


State and national mathematics achievement gaps at grade 4, 2007

Ten states had a smaller Black-White gap than the nation's 26-point gap in 2007 (Alaska, Delaware, DoDEA, Hawaii, Kentucky, Louisiana, Mississippi, Oklahoma, Texas, and West Virginia) and five had a gap that was larger (Connecticut, District of Columbia, Illinois, Nebraska, and Wisconsin). In 31 states, the gap was not significantly different from the nation's gap. Gaps that are different from the nation's gap are indicated with an asterisk (figure 9).

The fourth-grade mathematics gap in 2007 was statistically significant in all 46 states for which data could be reported. The gaps ranged from 14 points in Hawaii and West Virginia to 54 points in the District of Columbia.

Figure 9. The Black-White achievement score gap in mathematics for public school students at grade 4, by state or jurisdiction: 2007



* Significantly different (p<.05) from the nation (public) when comparing one state to the nation at a time.
¹ Department of Defense Education Activity (overseas and domestic schools).
 NOTE: States whose Black student population size was insufficient for comparison are omitted. Reporting standards not met for Idaho, Montana, North Dakota, Utah, Vermont, and Wyoming.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Mathematics Assessment.

Trends in state mathematics achievement gaps at grade 4, 1992–2007

The Black-White mathematics gap among the nation's public school fourth-graders was narrower in 2007 than in 1992, as Black students' scores showed a greater gain than White students' scores (figure 10, National results). From 2005 to 2007, there was no significant change in the gap.

In 35 states, both Black students and White students achieved higher average scores in mathematics from 1992 to 2007. Fifteen of these states also narrowed the achievement gap as Black students' scores increased more than White students' scores.

Short-term changes were also notable. In Illinois, New Jersey, and Virginia, average scores for both Black and White students increased between 2005 and 2007.

Narrowing of the Gap

In the following 15 states, the gap narrowed between **1992 and 2007** as gains of Black students outpaced the gains of White students.

- | | |
|-----------------------------|-----------------------|
| California | Michigan |
| Connecticut | Mississippi |
| Delaware | New Jersey |
| District of Columbia | Pennsylvania |
| Florida | South Carolina |
| Georgia | Texas |
| Louisiana | Virginia |
| Massachusetts | |

In **Rhode Island**, the gap narrowed between **2005 and 2007** as Black students' scores increased while those of White students did not change significantly.

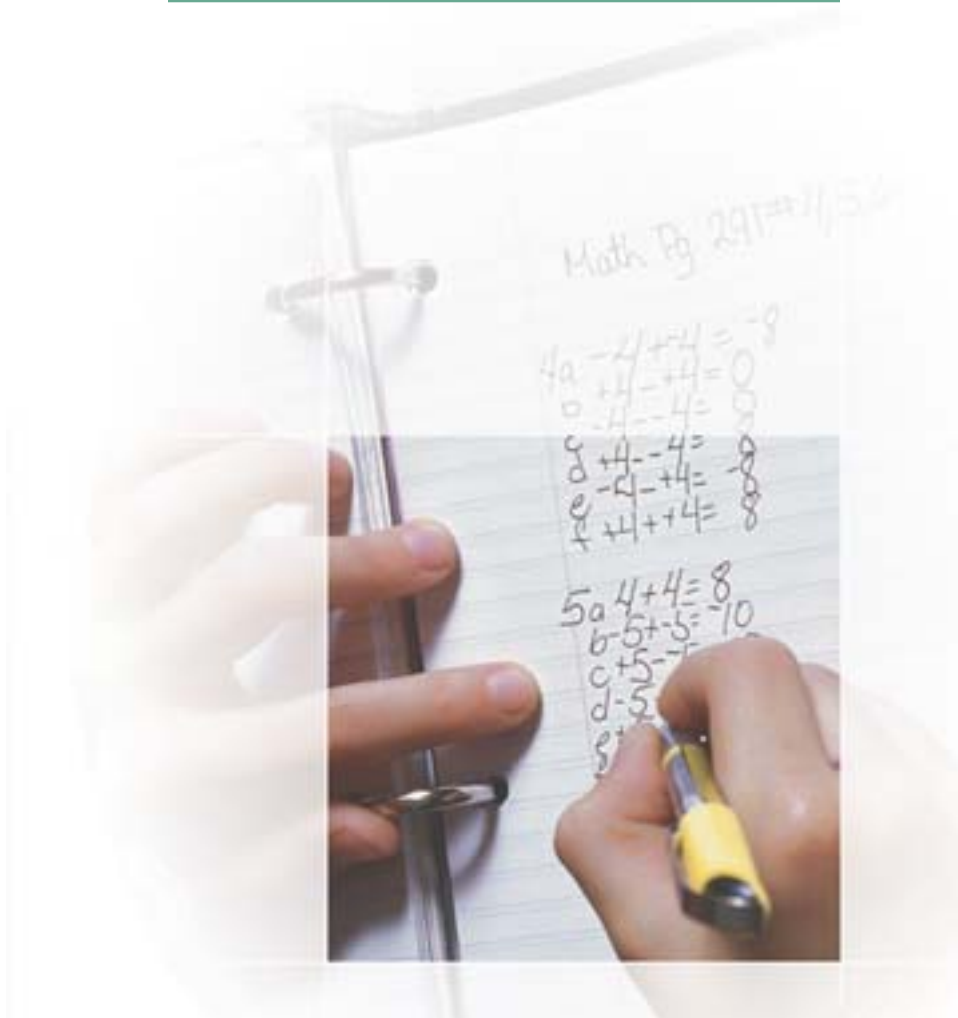
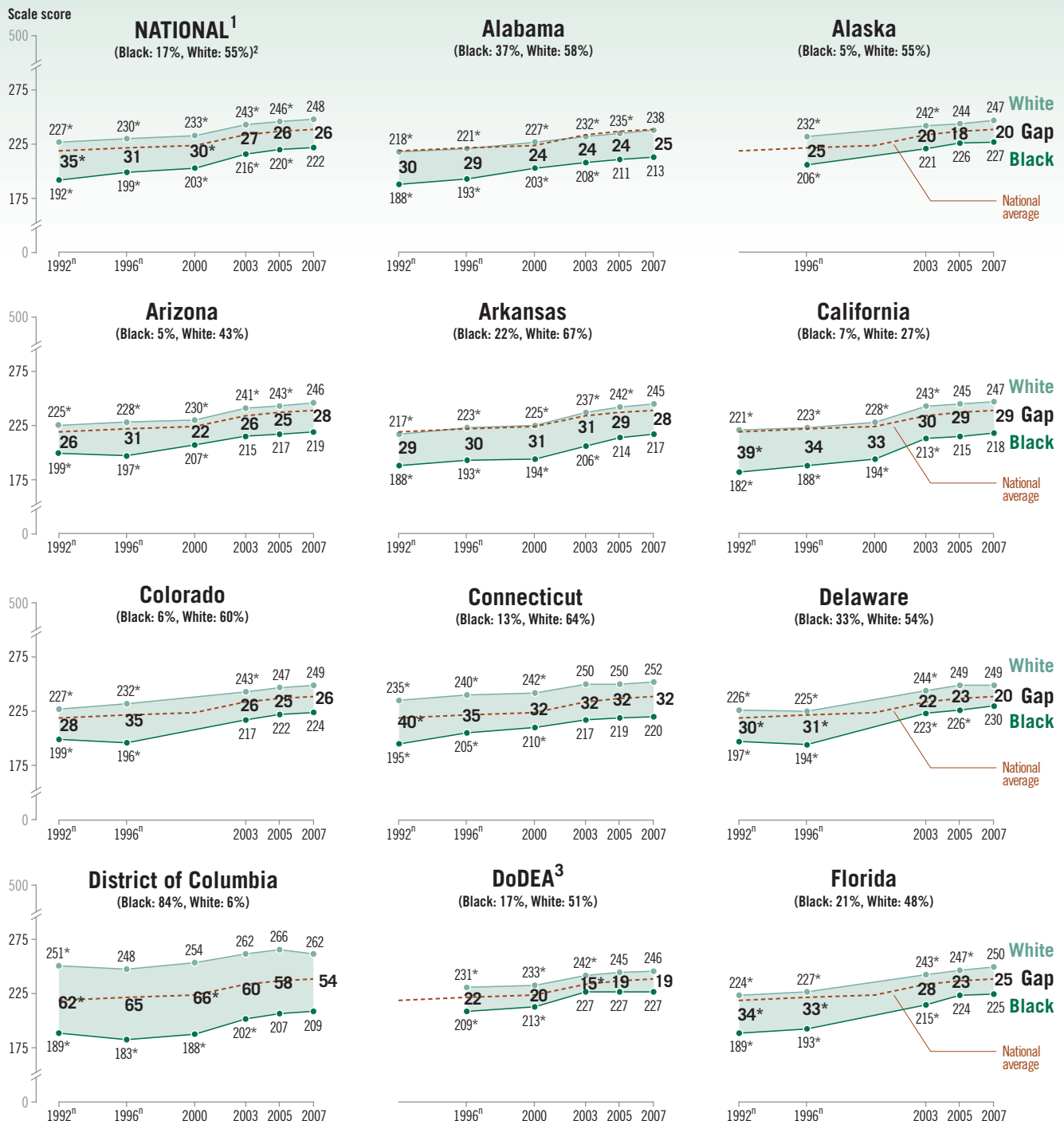
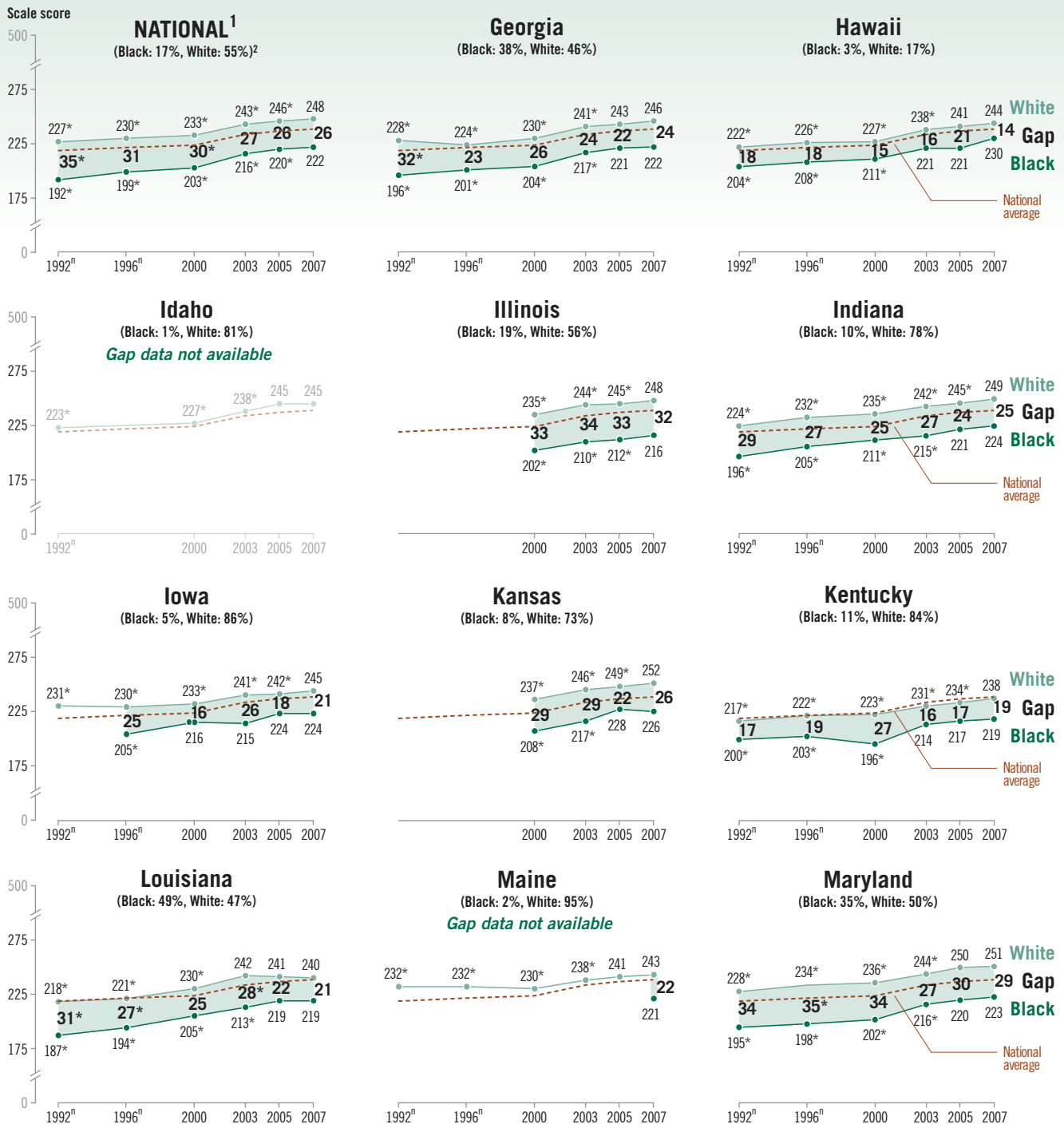


Figure 10. Gaps in average mathematics scores between Black and White public school students at grade 4, by state: Various years, 1992–2007



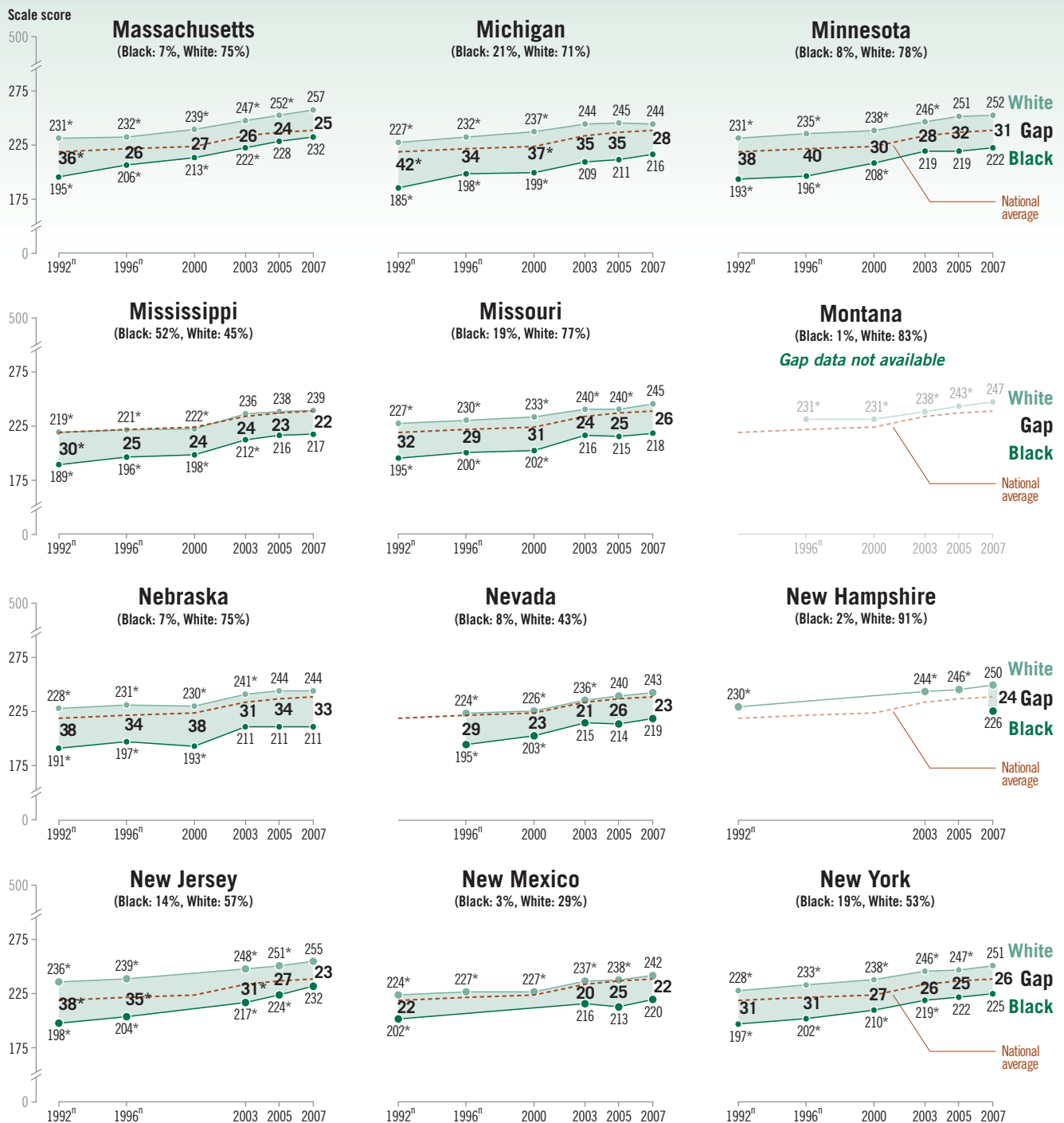
See notes at end of figure.

Figure 10. Gaps in average mathematics scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



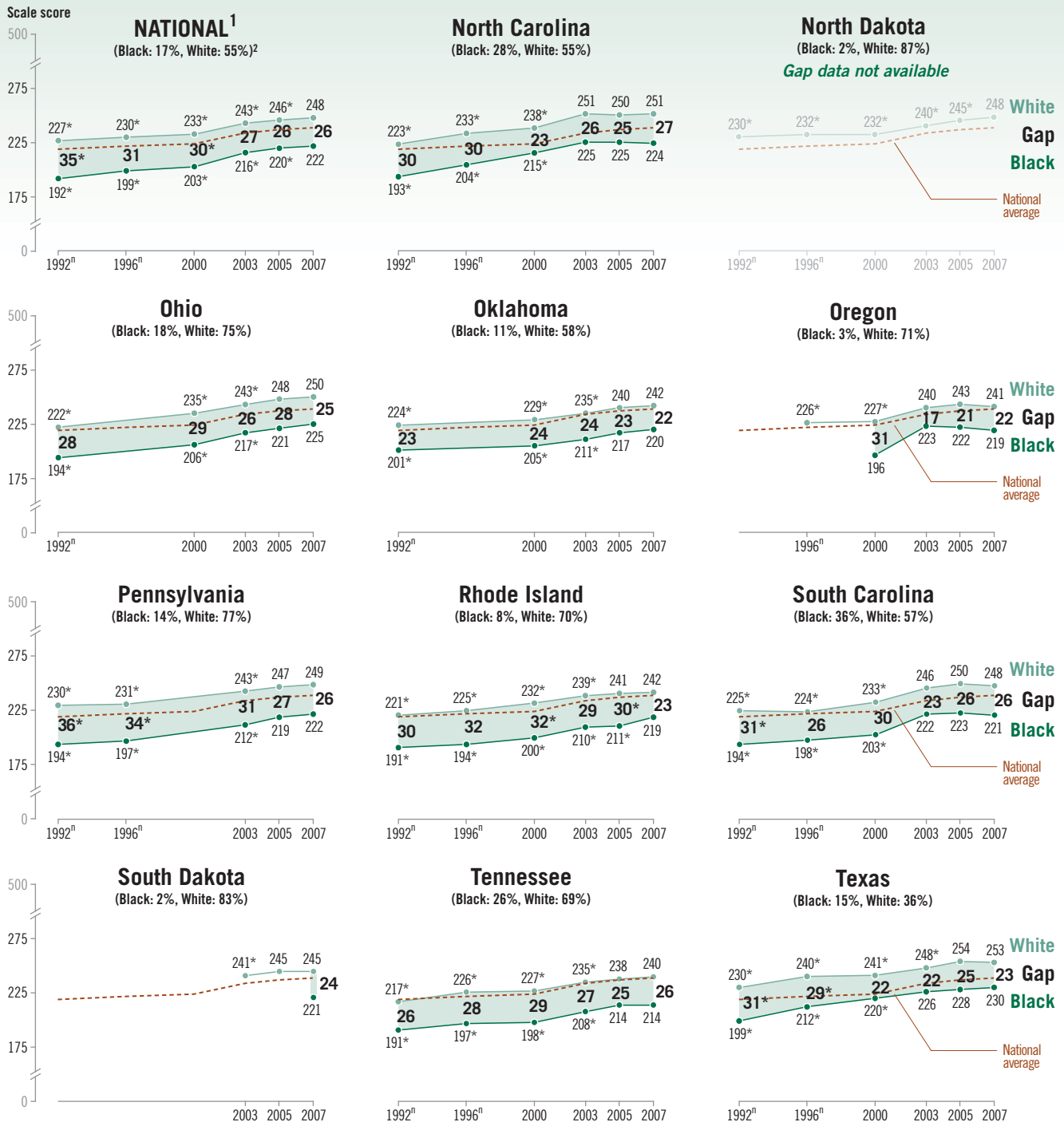
See notes at end of figure.

Figure 10. Gaps in average mathematics scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



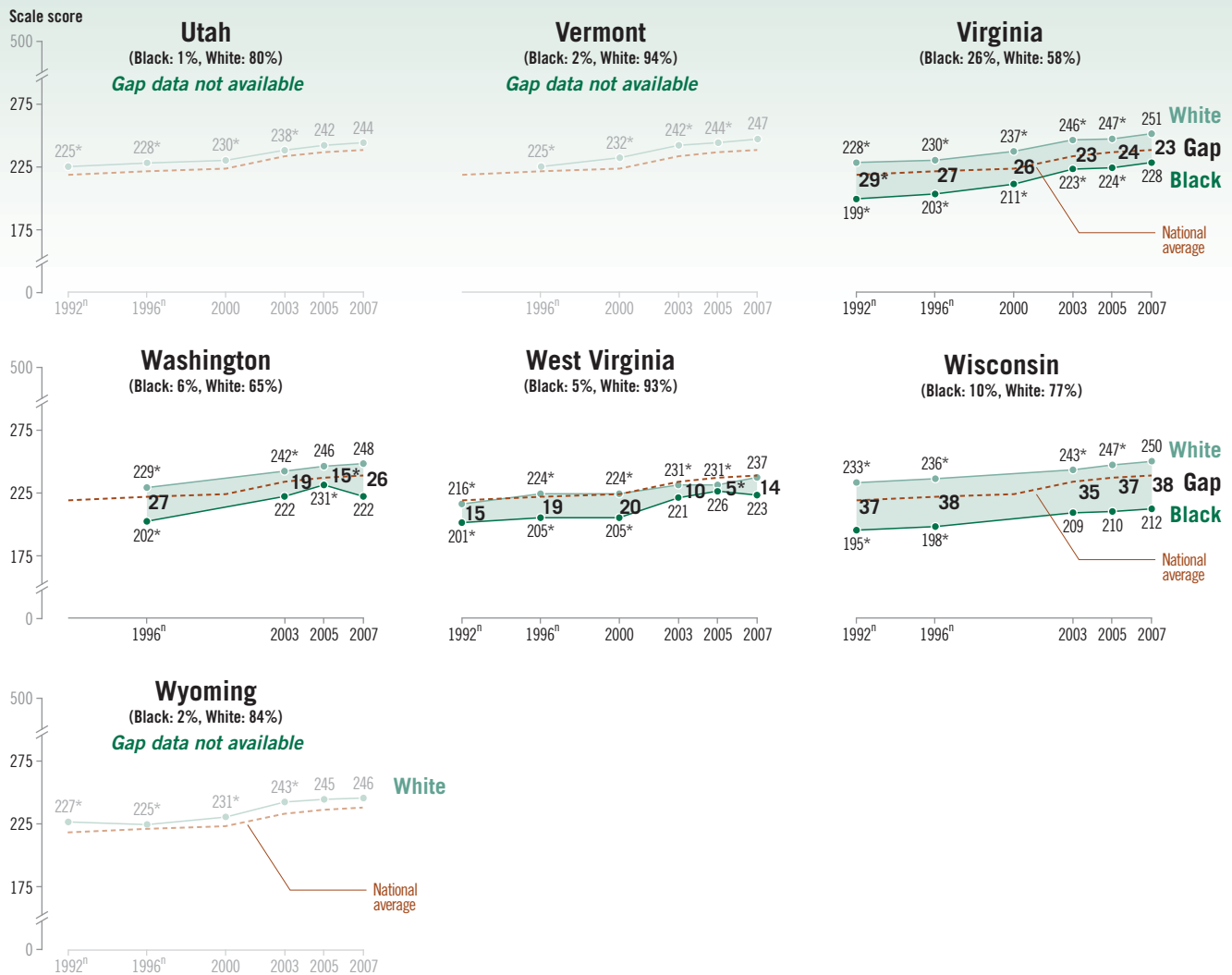
See notes at end of figure.

Figure 10. Gaps in average mathematics scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



See notes at end of figure.

Figure 10. Gaps in average mathematics scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

¹ National results for assessments prior to 2002 are based on the national sample, not on aggregated state samples.

² Black and White percentages are based on students tested in 2007.

³ Department of Defense Education Activity (overseas and domestic schools). Before 2005, DoDEA overseas and domestic schools were separate jurisdictions in NAEP. Pre-2005 data presented here were recalculated for comparability.

NOTE: Detail may not sum to totals due to rounding. Where data are not present, the jurisdiction did not participate or did not meet the minimum participation guidelines for reporting. State-level data were not collected in 1990. Comparative performance results may be affected by changes in exclusion rates for students with disabilities and English language learners in the NAEP samples.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Mathematics Assessments.

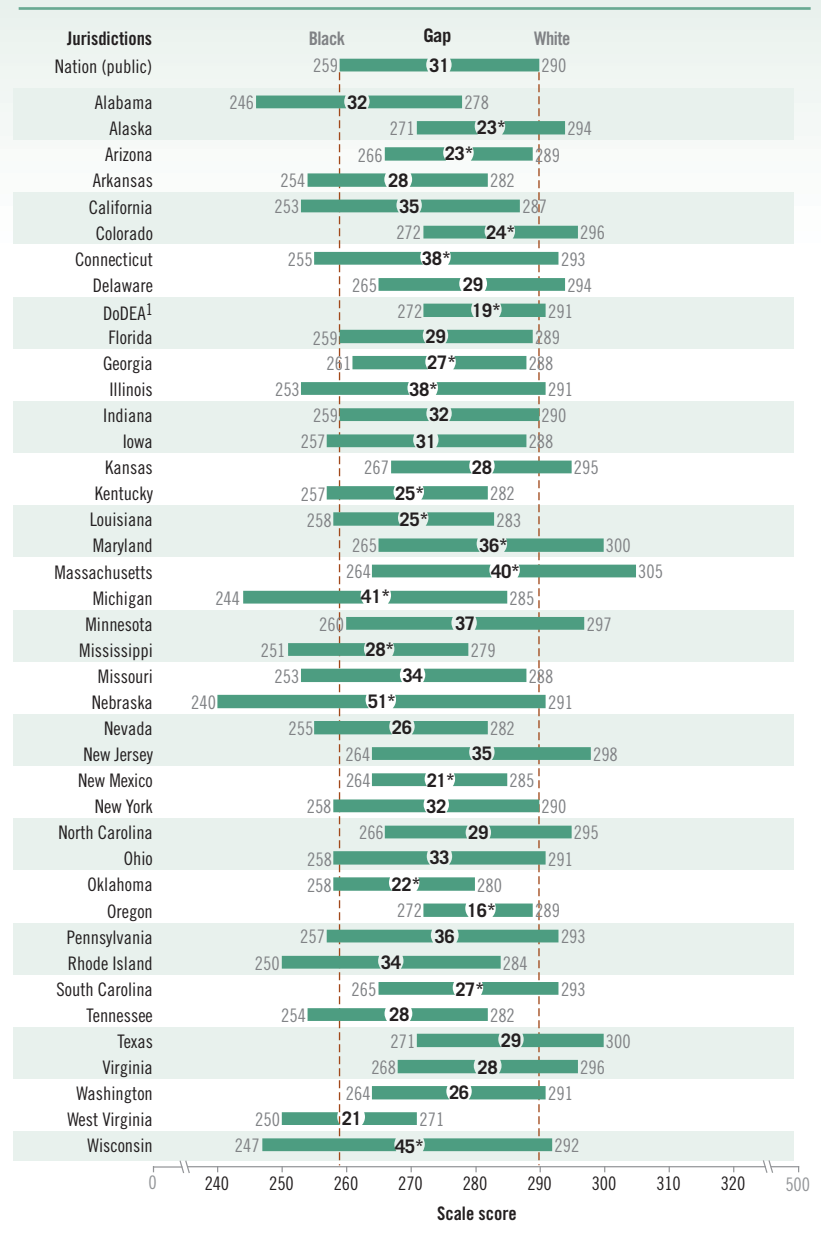


State and national mathematics achievement gaps at grade 8, 2007

Twelve states had a smaller gap than the nation's 31-point gap in 2007 (Alaska, Arizona, Colorado, DoDEA, Georgia, Kentucky, Louisiana, Mississippi, New Mexico, Oklahoma, Oregon, and South Carolina) and seven had a gap that was larger (Connecticut, Illinois, Maryland, Massachusetts, Michigan, Nebraska, and Wisconsin). In 22 states, the gap was not significantly different from the nation's gap. Gaps that are different from the nation's gap are indicated with an asterisk (figure 11).

The eighth-grade mathematics gap in 2007 was statistically significant in all 41 states for which data could be reported. The gaps ranged from 16 points in Oregon to 51 points in Nebraska.

Figure 11. The Black-White achievement score gap in mathematics for public school students at grade 8, by state or jurisdiction: 2007



* Significantly different (p<.05) from the nation (public) when comparing one state to the nation at a time.

¹ Department of Defense Education Activity (overseas and domestic schools).

NOTE: States whose Black or White student population size was insufficient for comparison are omitted. Reporting standards not met for District of Columbia, Hawaii, Idaho, Maine, Montana, New Hampshire, North Dakota, South Dakota, Utah, Vermont, and Wyoming.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Mathematics Assessment.

Trends in state mathematics achievement gaps at grade 8, 1990–2007

The national Black-White mathematics gap was not significantly narrower in 2007 than in 1990, despite higher average scores for both Black and White students in 2007 (figure 12, National results). The gap was narrower in 2007 than in 2005.

In 26 states, mathematics scores of both Black and White eighth-graders were higher in 2007 than in 1990. The 2007 gap was narrower in Arkansas, Colorado, Oklahoma, and Texas, as increases in Black students' scores were greater than those of their White peers.

Between 2005 and 2007, gaps narrowed in Arkansas and Florida as scores for Black eighth-graders increased while those of their White peers showed no change. In Colorado, scores for both groups increased, but a greater increase in Black students' scores caused the gap to narrow.

Narrowing of the Gap

In the following four states, the mathematics gap narrowed between **1990 and 2007** as gains of Black students outpaced the gains of White students.

Arkansas
Colorado

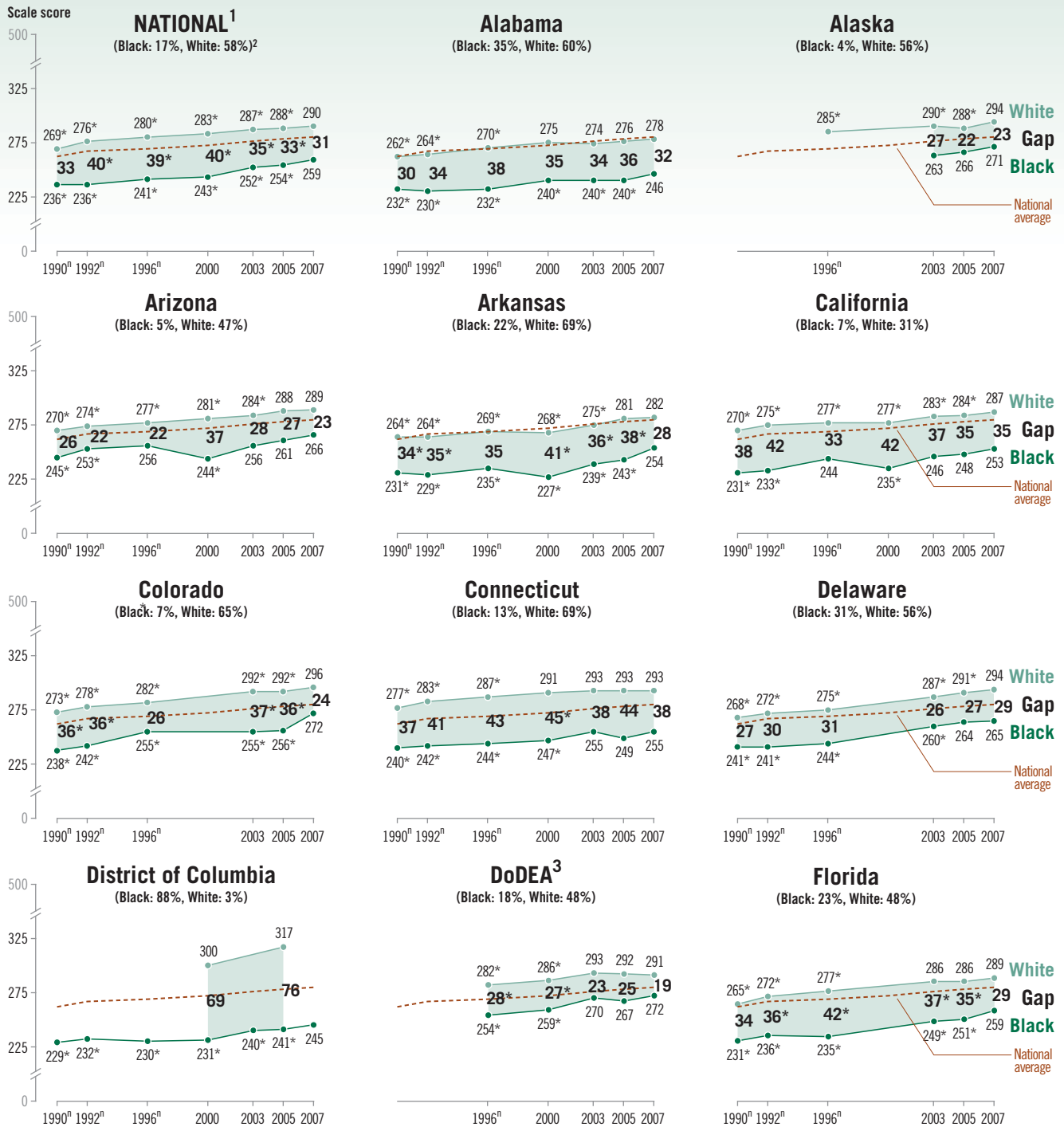
Oklahoma
Texas

In **Colorado**, the gap narrowed between **2005 and 2007** as Black students' scores showed greater increases than those of their White peers.

In **Arkansas** and **Florida**, the gap narrowed between **2005 and 2007** as Black students' scores increased while those of White students did not change significantly.

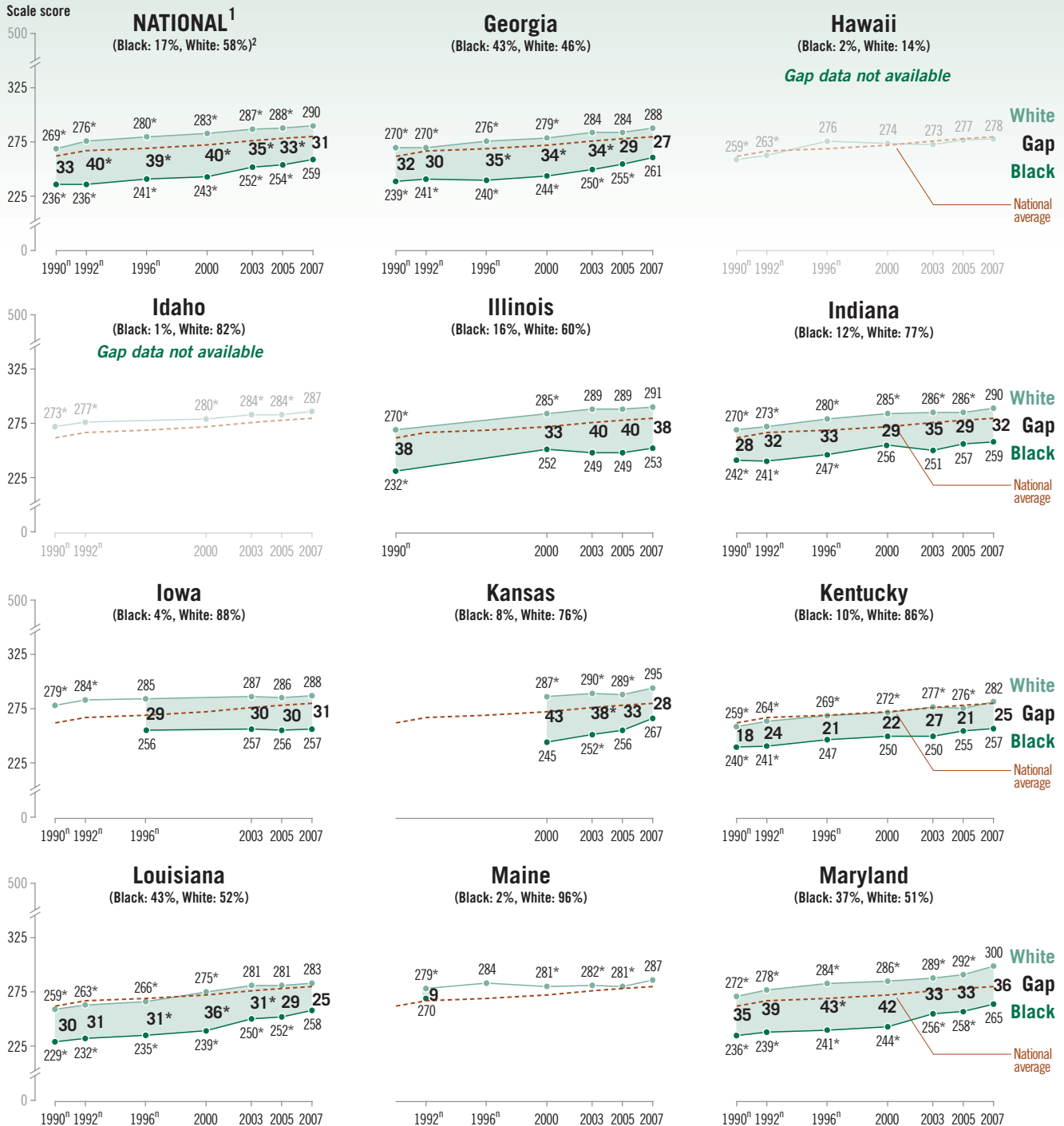


Figure 12. Gaps in average mathematics scores between Black and White public school students at grade 8, by state: Various years, 1990–2007



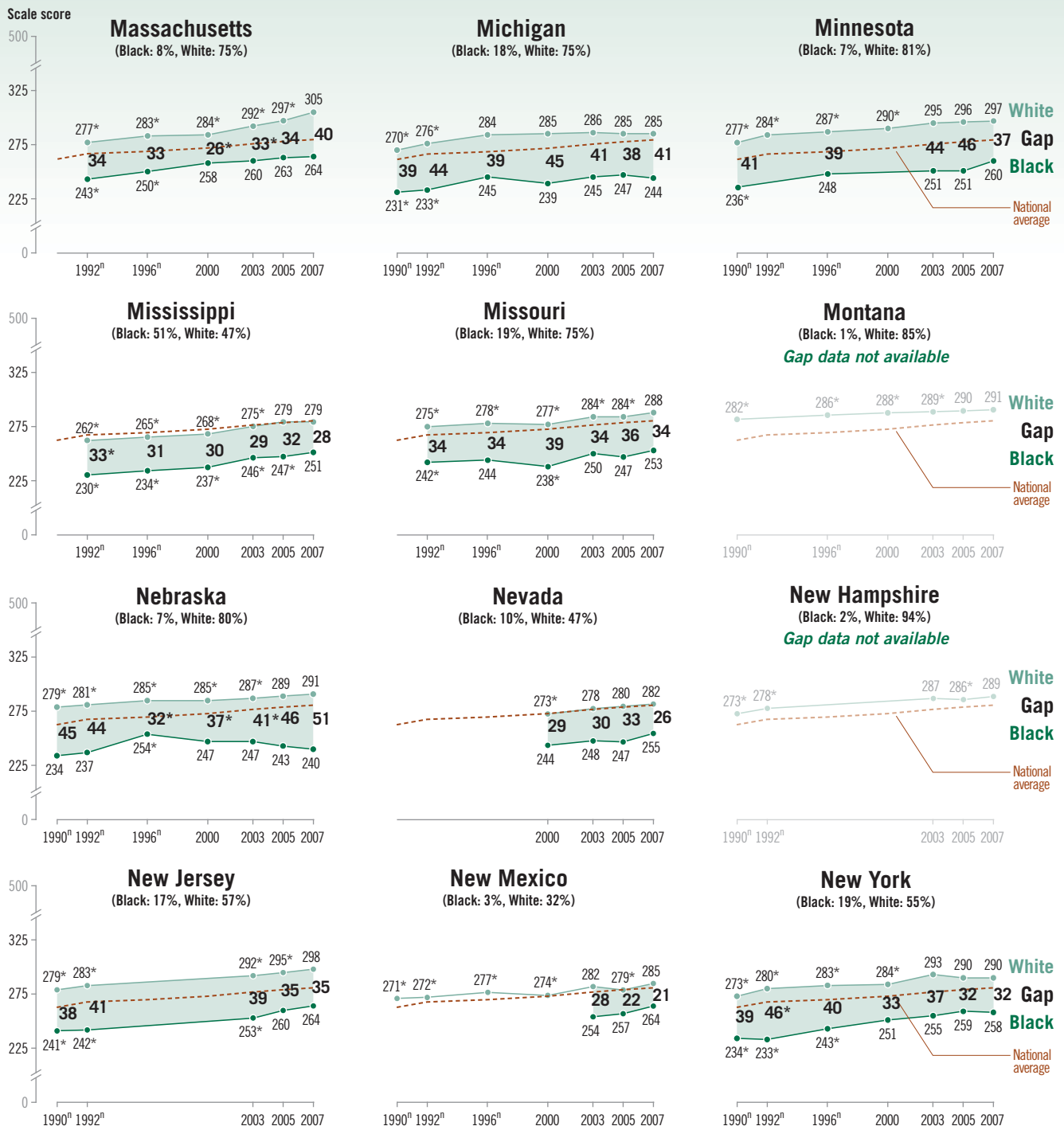
See notes at end of figure.

Figure 12. Gaps in average mathematics scores between Black and White public school students at grade 8, by state: Various years, 1990–2007—Continued



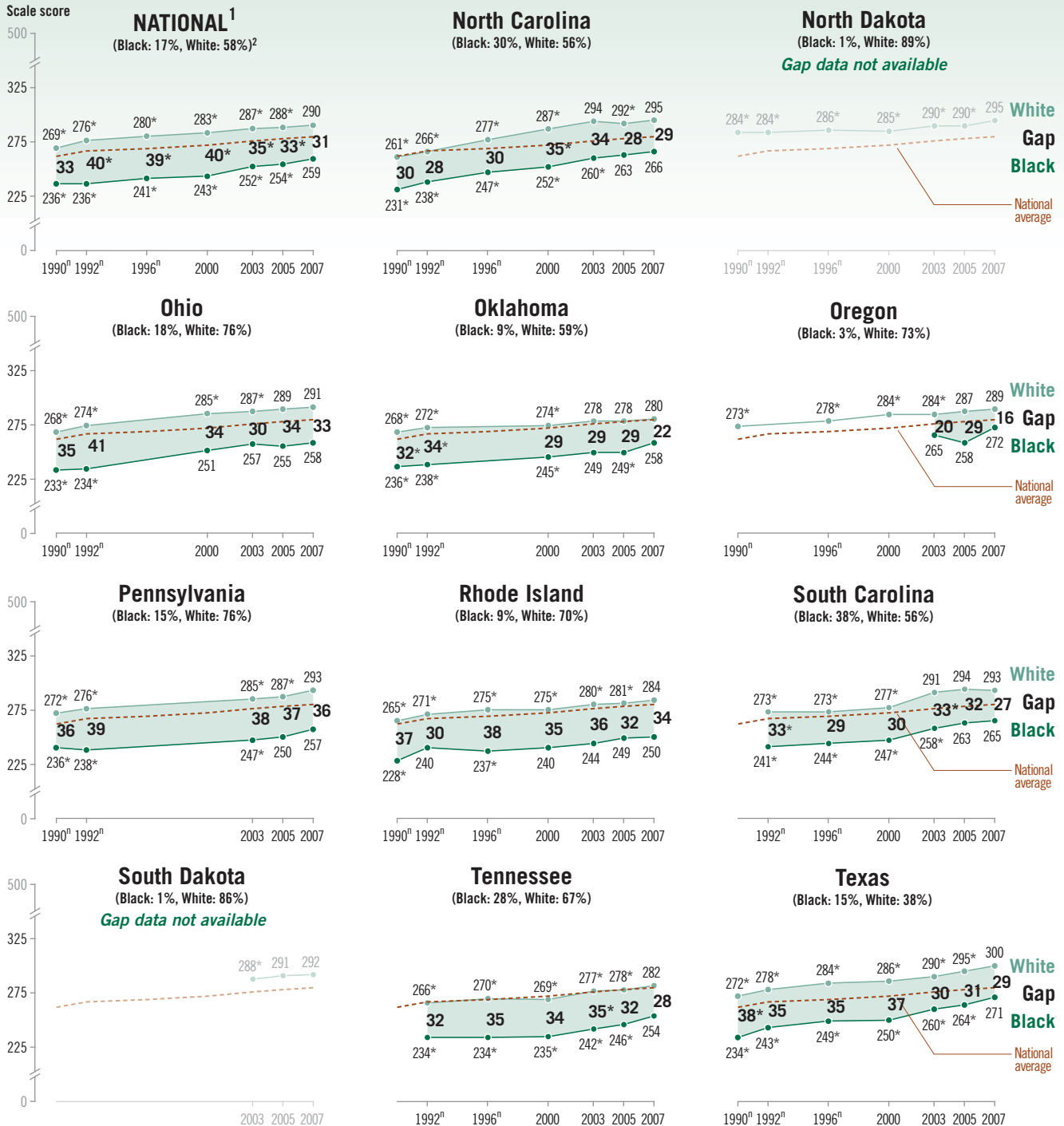
See notes at end of figure.

Figure 12. Gaps in average mathematics scores between Black and White public school students at grade 8, by state: Various years, 1990–2007—Continued



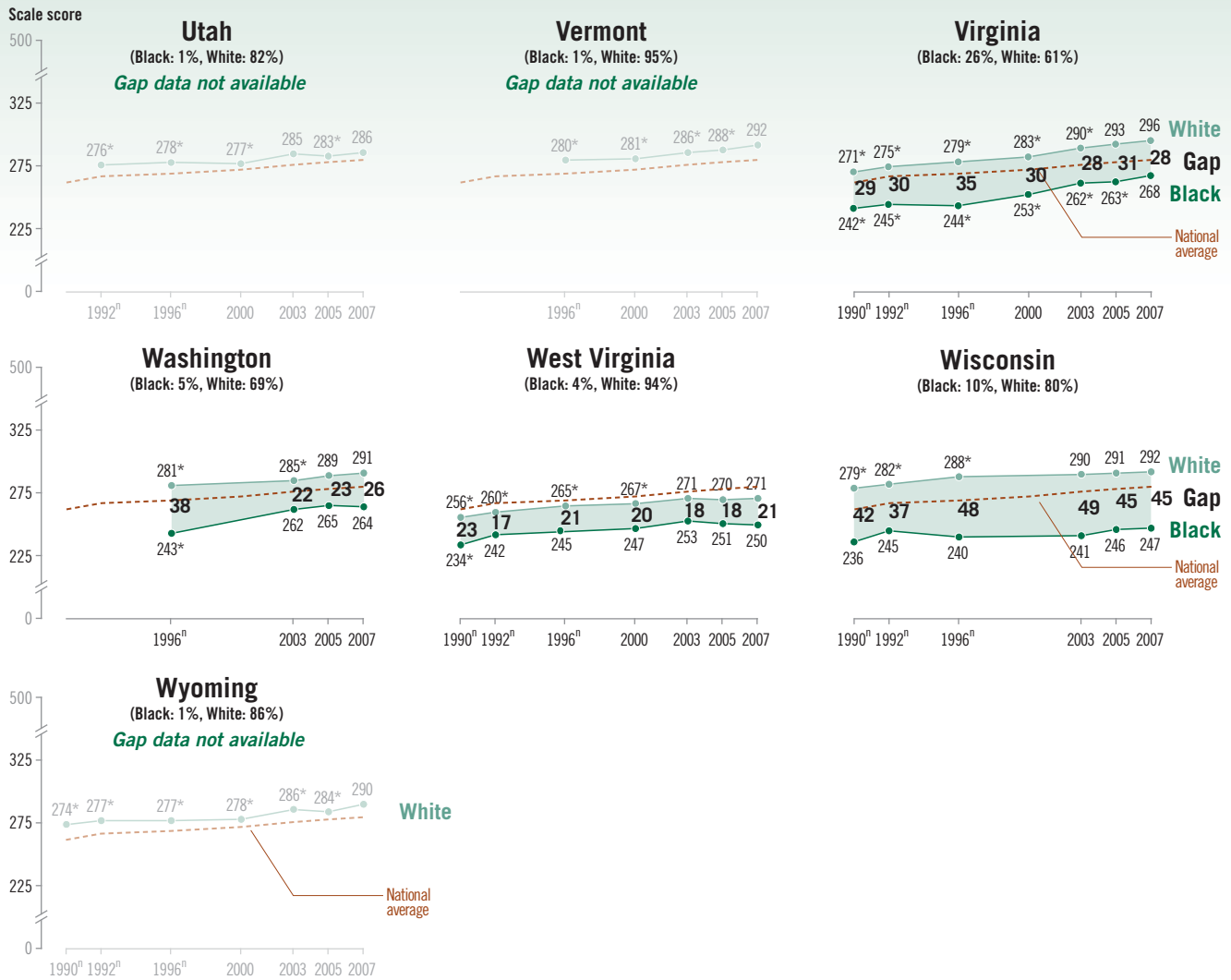
See notes at end of figure.

Figure 12. Gaps in average mathematics scores between Black and White public school students at grade 8, by state: Various years, 1990–2007—Continued



See notes at end of figure.

Figure 12. Gaps in average mathematics scores between Black and White public school students at grade 8, by state: Various years, 1990–2007—Continued



ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

¹ National results for assessments prior to 2002 are based on the national sample, not on aggregated state samples.

² Black and White percentages are based on students tested in 2007.

³ Department of Defense Education Activity (overseas and domestic schools). Before 2005, DoDEA overseas and domestic schools were separate jurisdictions in NAEP. Pre-2005 data presented here were recalculated for comparability.

NOTE: Detail may not sum to totals due to rounding. Where data are not present, the jurisdiction did not participate or did not meet the minimum participation guidelines for reporting. Comparative performance results may be affected by changes in exclusion rates for students with disabilities and English language learners in the NAEP samples.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1990–2007 Mathematics Assessments.

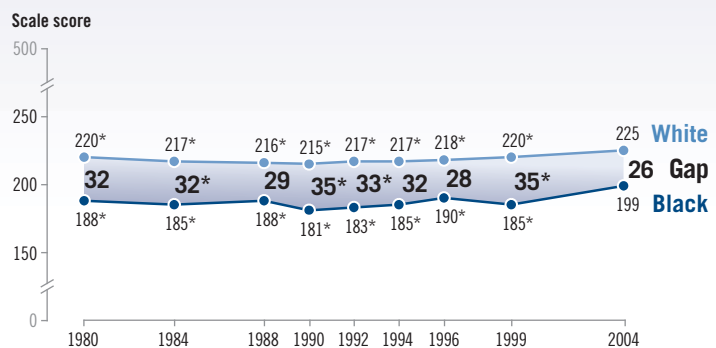
Long-Term Trend Results for Black and White 9- and 13-Year-Olds

Trends in reading scores and achievement gaps, 1980–2004

Reading scores for both Black and White 9-year-old students were higher in 2004 than on any previous long-term trend assessment (figure 13). The score gap in 2004 did not differ significantly from the gap in 1980, but was narrower than the gap in 1999, due to a greater increase in Black students' scores as compared to White students.

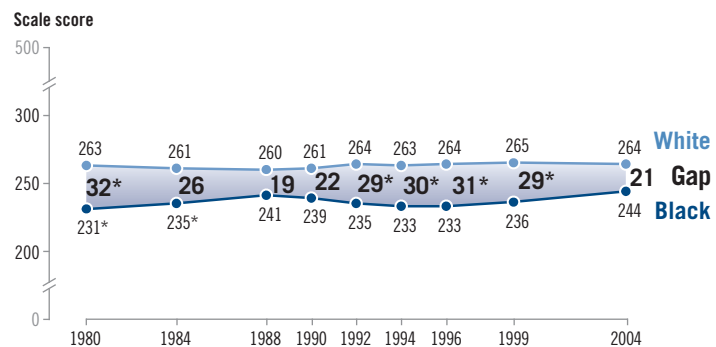
At age 13, reading scores for White students were not significantly different in 2004 than in 1980 (figure 14). For Black students, scores were higher in 2004 than in 1980, resulting in a narrowing of the gap. Scores did not change significantly for either Black or White students from 1999 to 2004, but the gap narrowed for that time period as well. A statistically significant change can occur over time in the gap between two scores even though the scores themselves do not change significantly because changes in gaps are calculated separately from changes in scores.

Figure 13. Trends in average reading scale scores and score gaps for White students and Black students at age 9: Various years, 1980–2004



* Significantly different ($p < .05$) from 2004.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1980–2004 Long-Term Trend Reading Assessments.

Figure 14. Trends in average reading scale scores and score gaps for White students and Black students at age 13: Various years, 1980–2004



* Significantly different ($p < .05$) from 2004.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1980–2004 Long-Term Trend Reading Assessments.

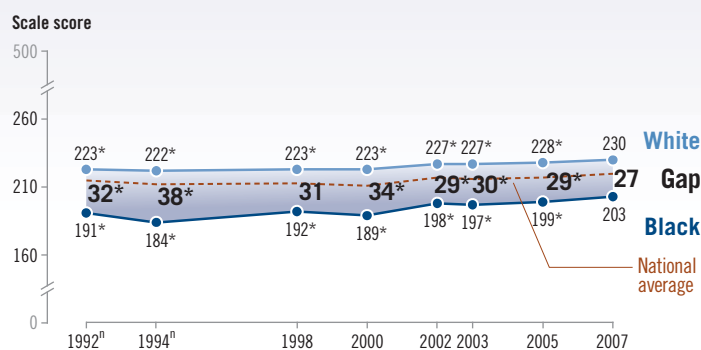
Main NAEP National Results for Black and White Fourth- and Eighth-Graders

Trends in reading scores and achievement gaps, 1992–2007

In main NAEP, the reading gap for Black and White fourth-graders narrowed in 2007 in comparison to both 1992 and 2005 (figure 15). Although scores for both Black and White students were higher in 2007 than in either comparison year, a greater increase in scores for Black students caused the gap to narrow. The 27-point gap in 2007 was narrower than in any previous assessment year except 1998.

Eighth-grade reading scores for both Black and White students were higher in 2007 than in either 1992 or 2005, but the gap in 2007 was not significantly different from either prior year (figure 16).

Figure 15. Reading achievement score gaps between Black and White public school students at grade 4: Various years, 1992–2007

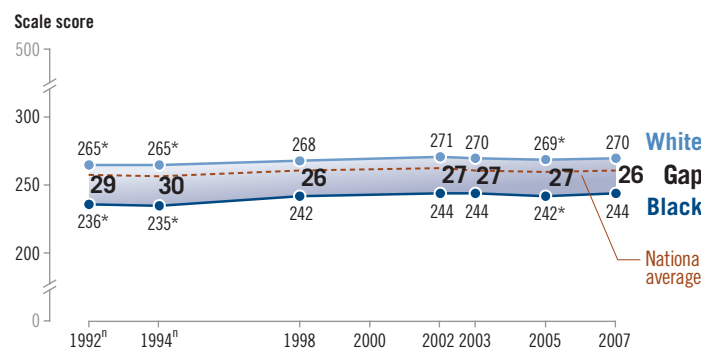


ⁿ Accommodations were not permitted for this assessment.

* Significantly different (p<.05) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Reading Assessments.

Figure 16. Reading achievement score gaps between Black and White public school students at grade 8: Various years, 1992–2007



ⁿ Accommodations were not permitted for this assessment.

* Significantly different (p<.05) from 2007.

NOTE: Data were not collected at grade 8 in 2000.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Reading Assessments.

Reading scores and achievement gaps by gender, 1992–2007

Average reading scores were higher in 2007 than in 1992 for Black and for White fourth-graders, regardless of gender (figure 17). Among males, the gap narrowed as the scores of Black males increased more than those of their White peers. Among fourth-grade females, the Black-White gap did not change significantly.

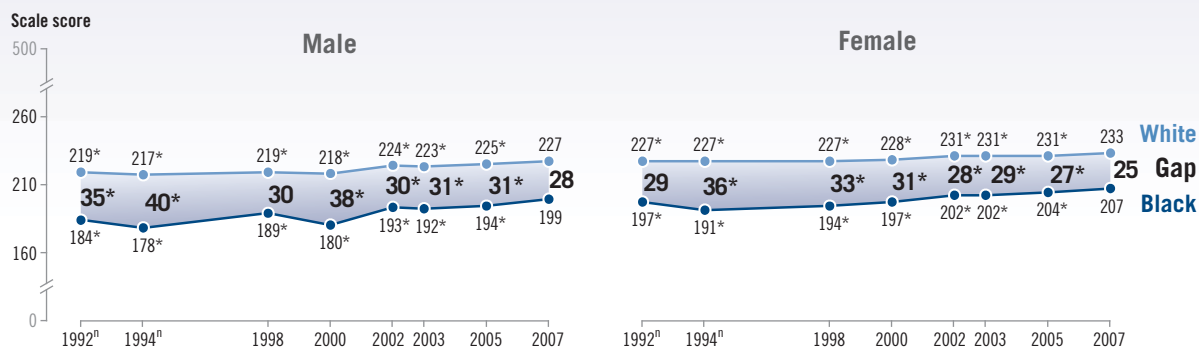
Fourth-grade reading scores were higher in 2007 than in 2005 for both Black and White males and females, and the achievement gaps narrowed for both groups during this period, as the scores of Black fourth-graders increased more than those of their White peers.

Average reading scores for eighth-graders were higher in 2007 than in 1992 for Black and for White students, regardless of gender (figure 18). However, the 2007 gaps in eighth-grade reading achievement showed no significant differences from the 1992 gaps for either males or females.

From 2005 to 2007, average reading scores for eighth-graders increased for both Black and White males. Scores increased for Black females but not for White females. However, the Black-White gap did not change significantly for either gender during this period.



Figure 17. Gaps in average reading scores between Black and White public school students at grade 4, by gender: Various years, 1992–2007

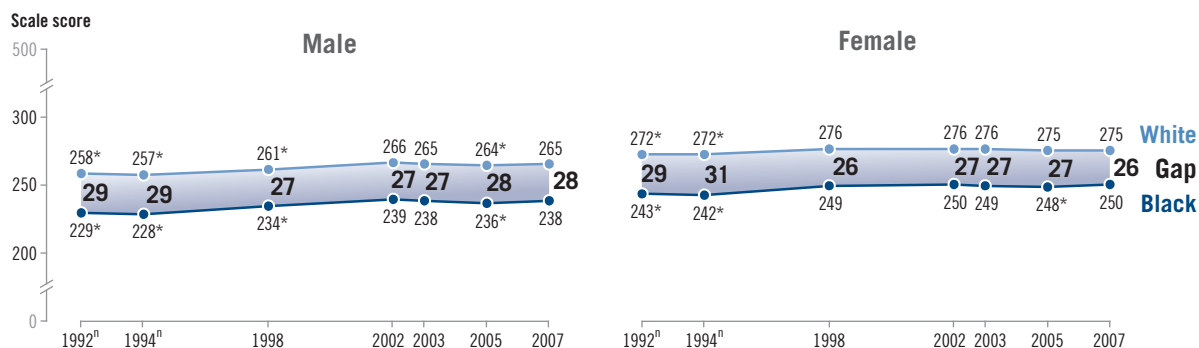


ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Reading Assessments.

Figure 18. Gaps in average reading scores between Black and White public school students at grade 8, by gender: Various years, 1992–2007



ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

NOTE: Data were not collected at grade 8 in 2000.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Reading Assessments.

Reading scores and achievement gaps by family income, 2003–2007

NAEP uses student eligibility for free or reduced-price school lunch as an indicator of family income. At grade 4, reading scores were higher in 2007 than in 2003 for both Black and White public school students, regardless of school-lunch eligibility (figure 19). The gap in 2007 for not-eligible students was narrower than in 2003, while the gap for students eligible for free lunch was narrower than in either previous assessment.

At grade 8, scores were higher for Black and White not-eligible students only, comparing 2007 with 2005 (figure 20). There were no statistically significant changes in the sizes of the gaps.

Table 2. Percentage of public school students assessed in NAEP reading by eligibility for free or reduced-price school lunch, race/ethnicity and grade: 2003, 2005, and 2007

	Not eligible		Eligible for reduced-price lunch		Eligible for free lunch	
	Black	White	Black	White	Black	White
Grade 4						
2007	26	73	7	6	66	21
2005	25	72	8	7	66	20
2003	24	72	9	8	65	18
Grade 8						
2007	32	76	7	5	59	18
2005	32	75	9	6	57	17
2003	32	76	9	6	56	14

NOTE: Detail may not sum to totals due to rounding.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003, 2005, and 2007 Reading Assessments.

Eligibility for free and reduced-price lunch

NAEP collects data on students' eligibility for the National School Lunch Program (NSLP)—sometimes referred to as the free and reduced-price school lunch program—as an indicator of family economic status. Eligibility for free and reduced-price lunches is based on students' family income in relation to the federally established poverty level.

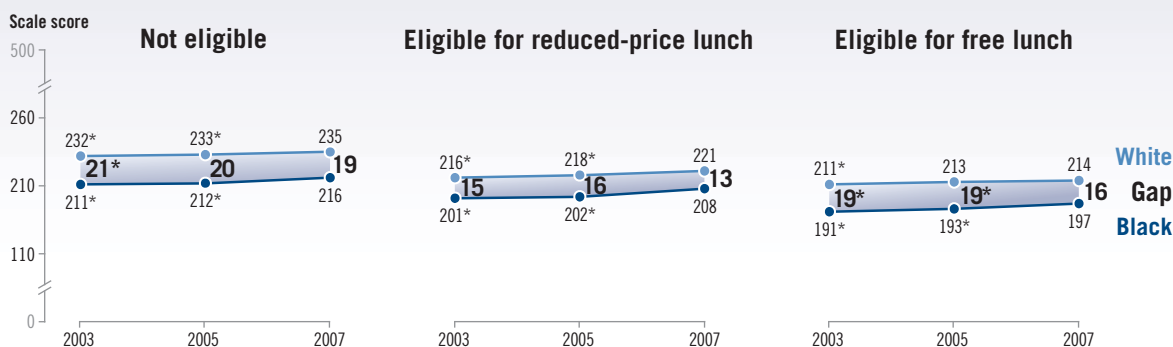
Not eligible: Students who are not eligible for the program because their family's income is above 185 percent of the poverty level.

Eligible for reduced-price lunch: Students who are eligible for reduced-price lunch because their family's income is between 130 percent and 185 percent of the poverty level.

Eligible for free lunch: Students who are eligible for free lunch because their family's income is below 130 percent of the poverty level.

As a result of improvements in the quality of the data on students' eligibility for NSLP, the percentage of students for whom information was not available has decreased in comparison to the percentages reported prior to the 2003 assessment. Therefore, trend comparisons are only made back to 2003 in this report.

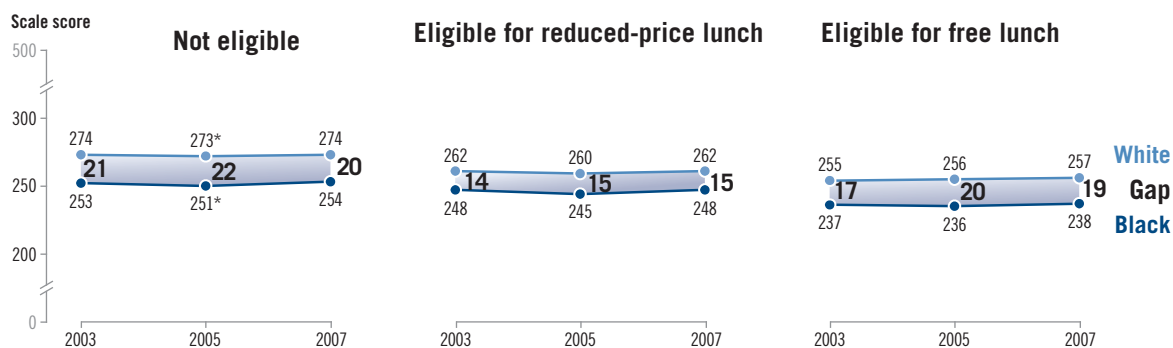
Figure 19. Gaps in average reading scores between Black and White public school students at grade 4, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007



* Significantly different (p<.05) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003, 2005, and 2007 Reading Assessments.

Figure 20. Gaps in average reading scores between Black and White public school students at grade 8, by eligibility for free or reduced-price school lunch: 2003, 2005, and 2007



* Significantly different (p<.05) from 2007.

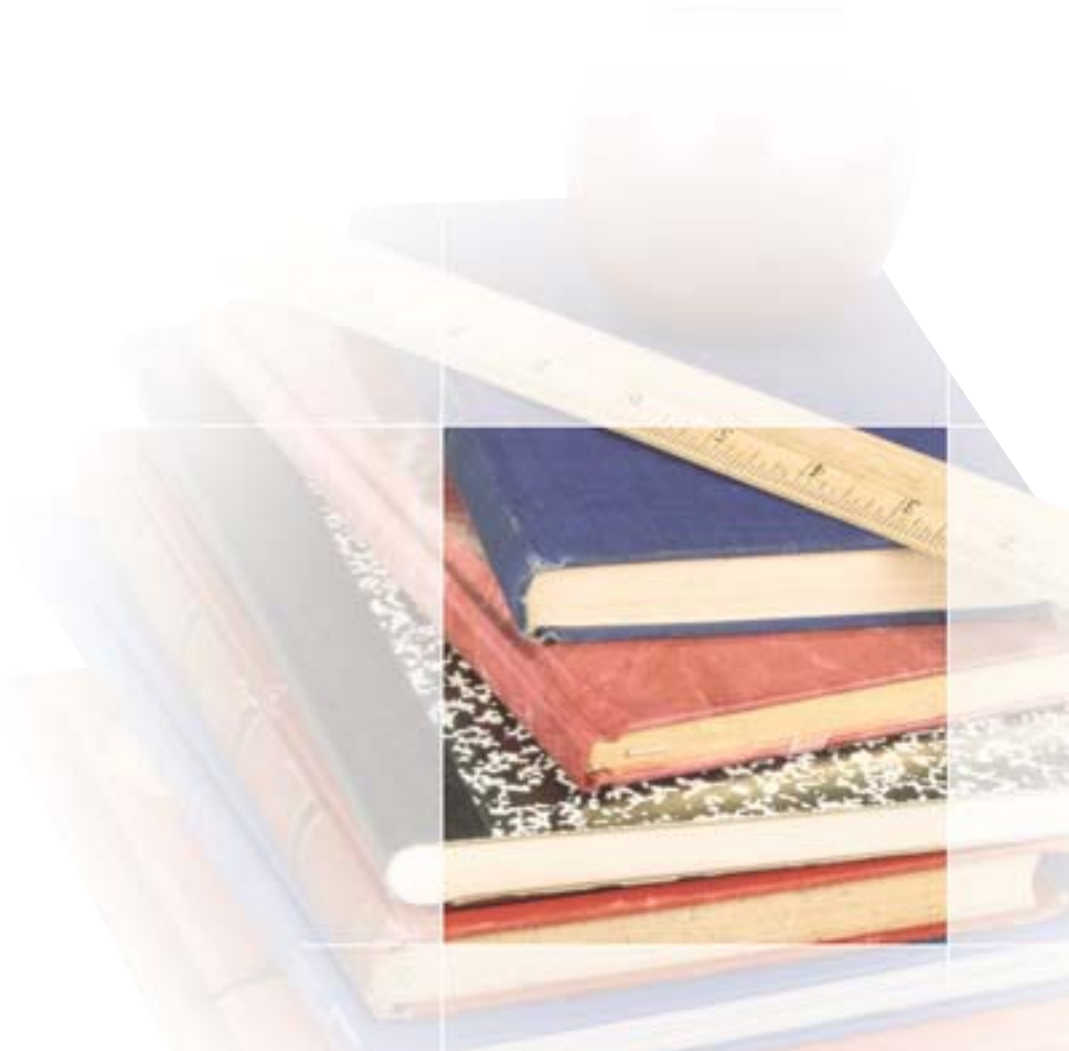
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003, 2005, and 2007 Reading Assessments.

Main NAEP State Results for Black and White Fourth- and Eighth-Graders

The NAEP state reading assessments were administered to public school students in fourth grade in 1992, 1994, 1998, 2002, 2003, 2005, and 2007 and in eighth grade in 1998, 2002, 2003, 2005, and 2007. Before 2003, states were not required to participate in NAEP in order to qualify for Title I education funds. Typically, 40 or more states participated in each assessment prior to 2003. In 2003, 2005, and 2007, all 50 states, the District of Columbia, and the DoDEA schools participated.

State results are presented in two ways. Comparisons of fourth-grade reading gaps in 2007 between each state and the nation are presented in figure 21.

Comparisons of the reading gaps within a state over time are presented in a series of small graphs in figure 22. At the top left of each two-page spread, the reading scores and gaps for the nation are presented for reference. Each state figure, as well as the national figure, also contains a dotted red line representing the national average for public school students. The data for the national averages are located in appendix B in table B-4.

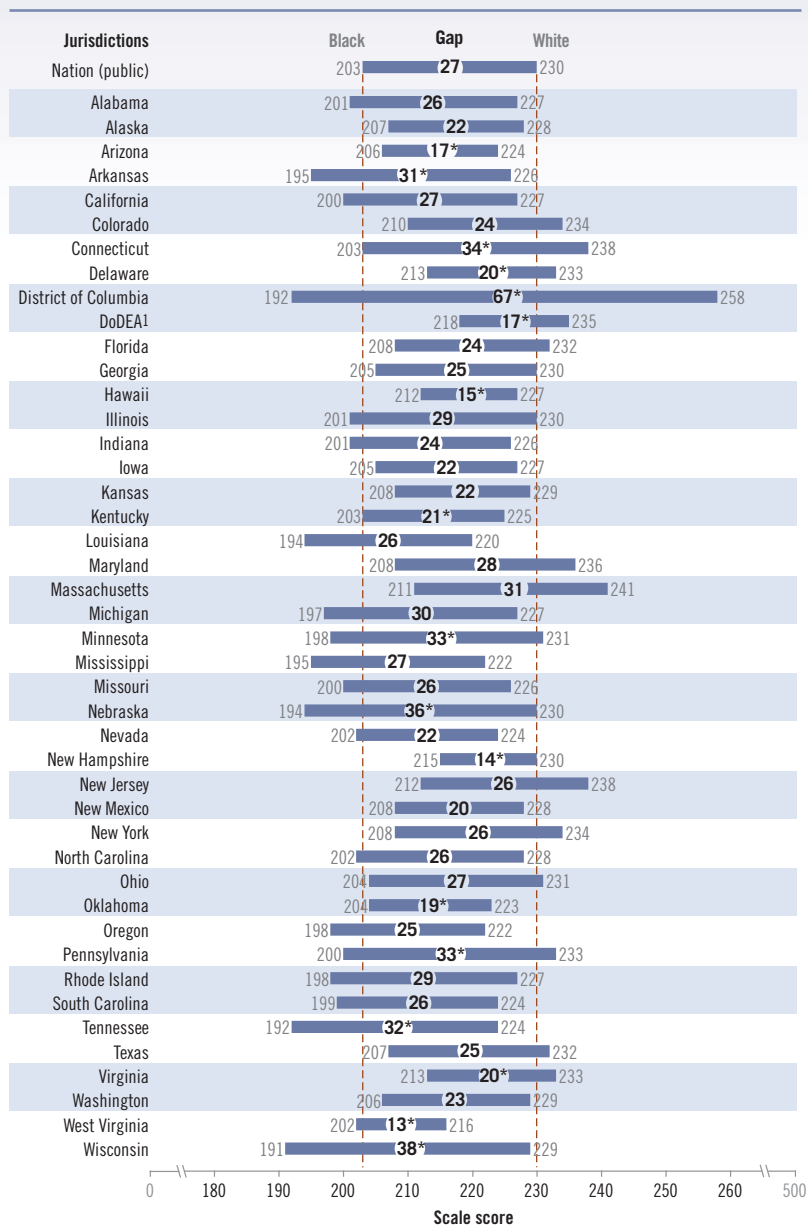


State and national reading achievement gaps at grade 4, 2007

Nine states had a Black-White gap that was smaller than the nation's 27-point gap in 2007 (Arizona, Delaware, DoDEA, Hawaii, Kentucky, New Hampshire, Oklahoma, Virginia, and West Virginia) and eight had a gap that was larger (Arkansas, Connecticut, District of Columbia, Minnesota, Nebraska, Pennsylvania, Tennessee, and Wisconsin). In 27 states, the gap was not different from the national gap. Gaps that are different from the national gap are indicated with an asterisk (figure 21).

The Black-White grade 4 reading gap in 2007 was significant in all 44 states for which data could be reported. The gaps ranged from 13 points in West Virginia to 67 points in the District of Columbia.

Figure 21. The Black-White achievement score gap in reading for public school students at grade 4, by state or jurisdiction: 2007



* Significantly different ($p < .05$) from the nation (public) when comparing one state to the nation at a time.
¹ Department of Defense Education Activity (overseas and domestic schools).
 NOTE: States whose Black student population size was insufficient for comparison are omitted. Reporting standards not met for Idaho, Maine, Montana, North Dakota, South Dakota, Utah, Vermont, and Wyoming.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Reading Assessment.

Trends in state reading achievement gaps at grade 4, 1992–2007

The Black-White reading gap among the nation’s public school fourth-graders was narrower in 2007 than in 1992 as average scores for Black students demonstrated a larger increase than average scores for White students (figure 22, National results).

In 13 states, both Black and White fourth-graders achieved higher average scores in reading during this period. In three states—Delaware, Florida, and New Jersey—the gap was narrower in 2007 than in 1992 as Black students’ scores increased more than those of White students.

In addition, gaps narrowed from 2005 to 2007 in Alabama, Arizona, and Virginia.

Narrowing of the Gap

In the following three states, the reading gap was narrower in **2007 than in 1992**, as Black students’ scores increased more than those of their White peers.

Delaware
Florida

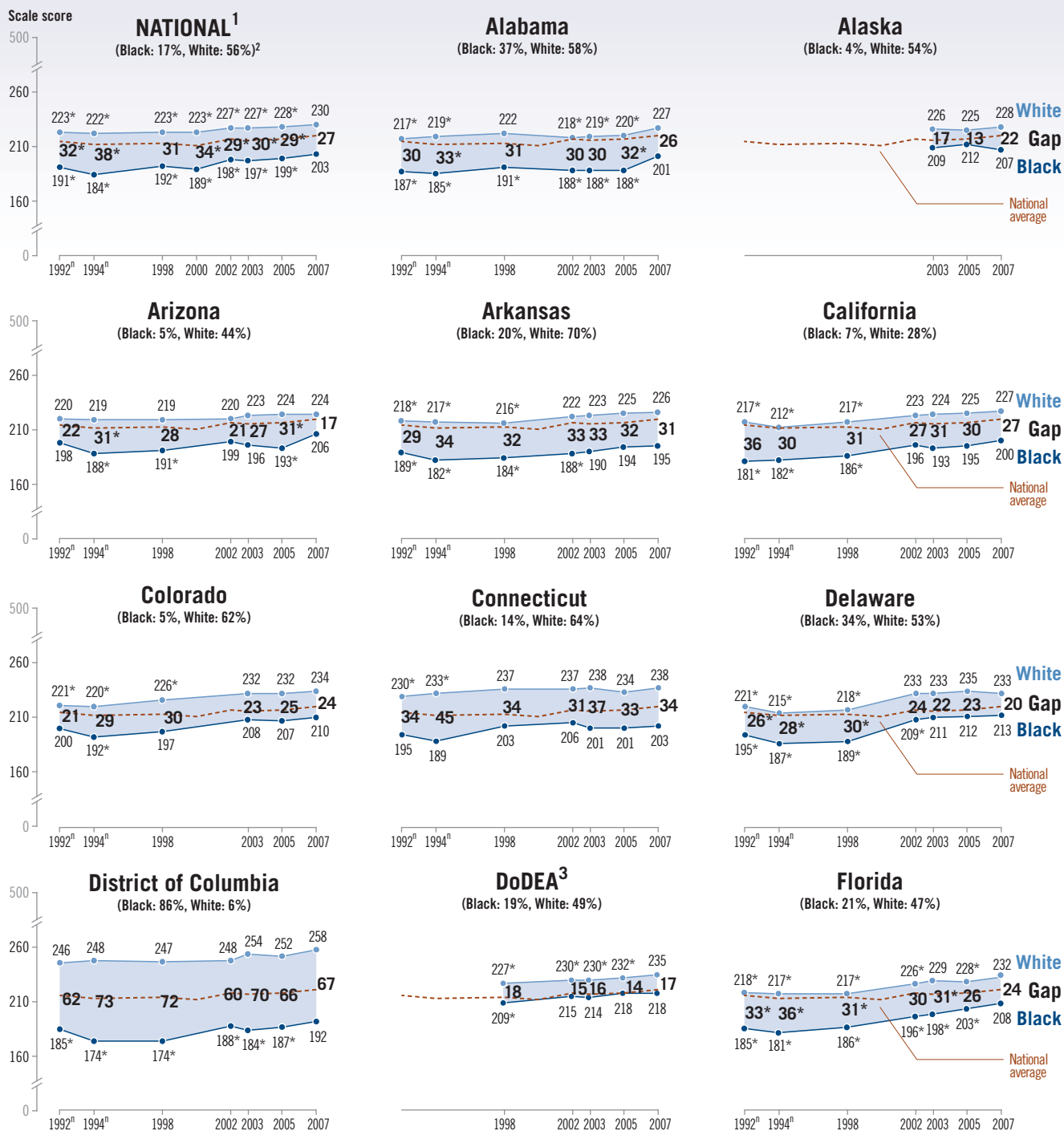
New Jersey

In **Alabama**, the reading gap narrowed between **2005 and 2007** as Black students’ scores increased more than White students’ scores.

In **Arizona and Virginia**, the reading gap narrowed between **2005 and 2007** as Black students’ scores increased while those of White students did not change significantly.

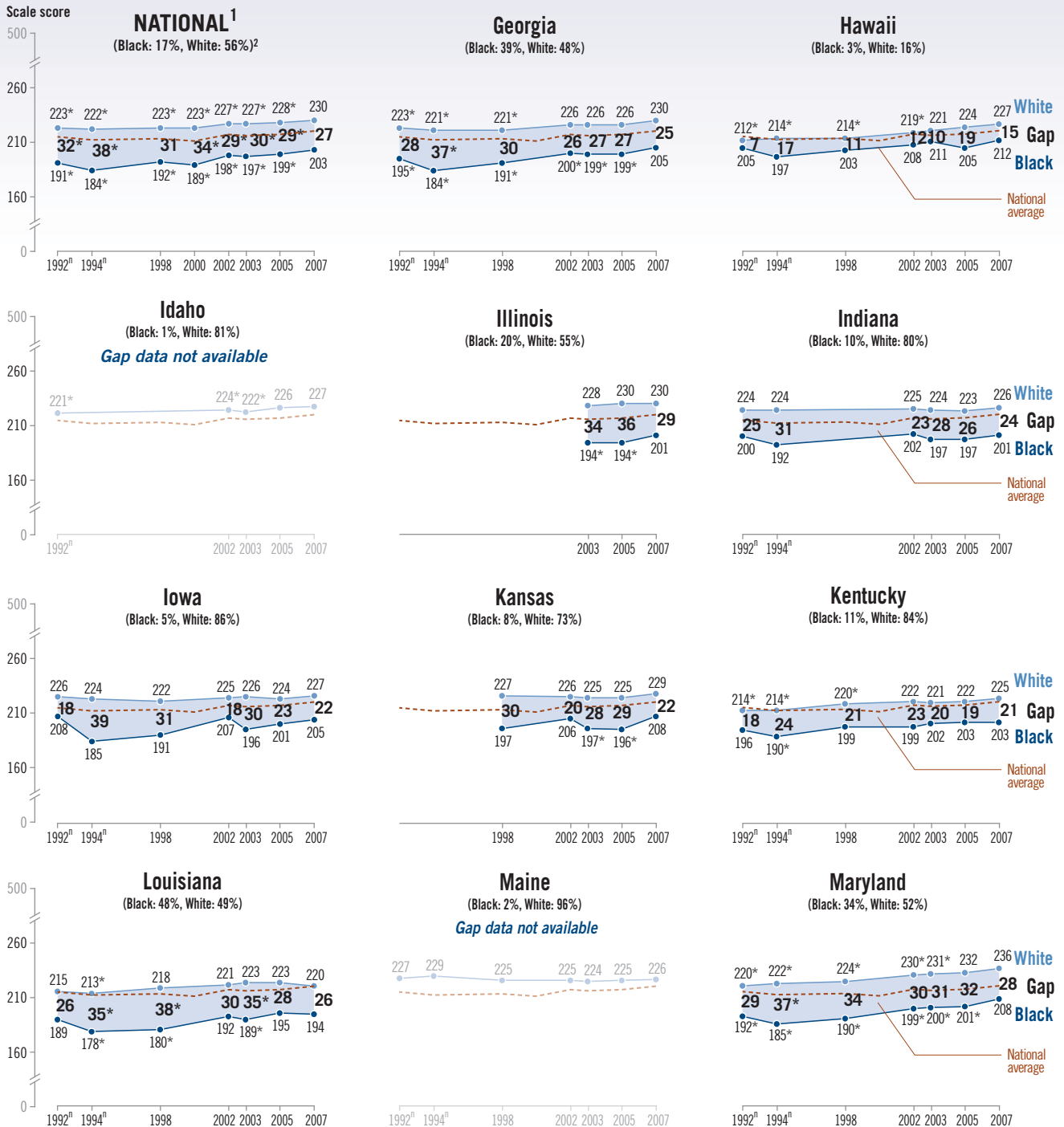


Figure 22. Gaps in average reading scores between Black and White public school students at grade 4, by state: Various years, 1992–2007



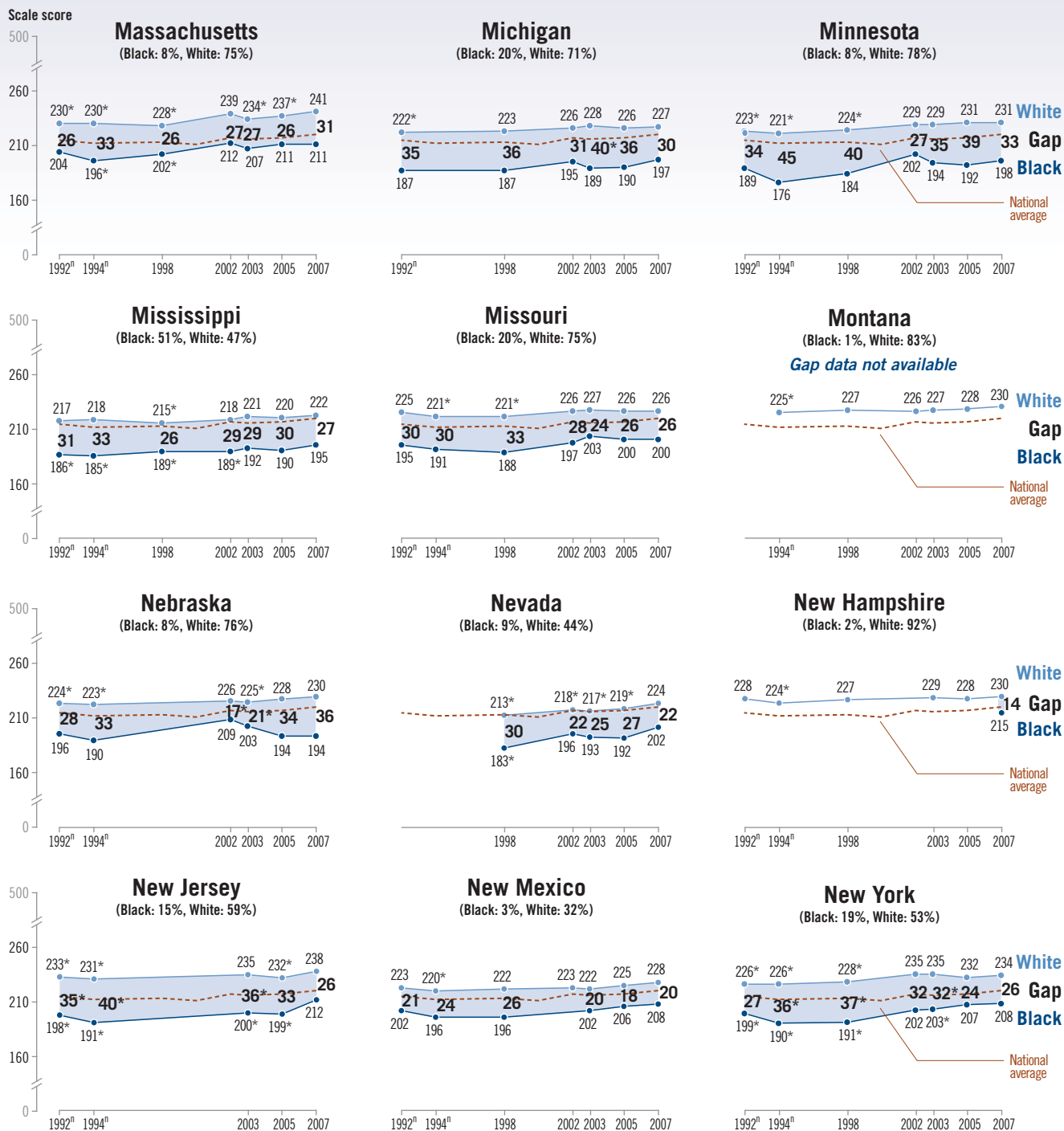
See notes at end of figure.

Figure 22. Gaps in average reading scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



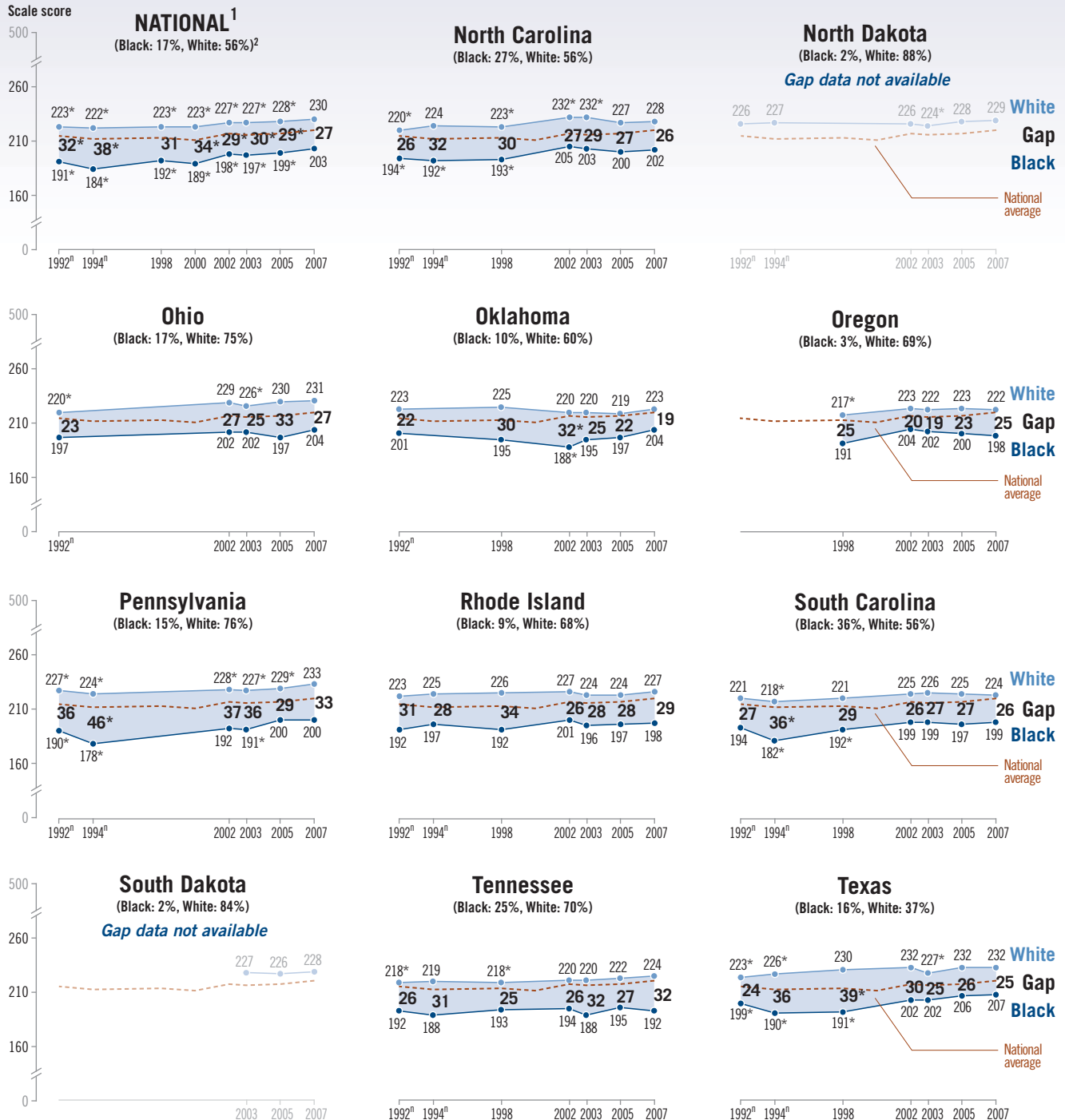
See notes at end of figure.

Figure 22. Gaps in average reading scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



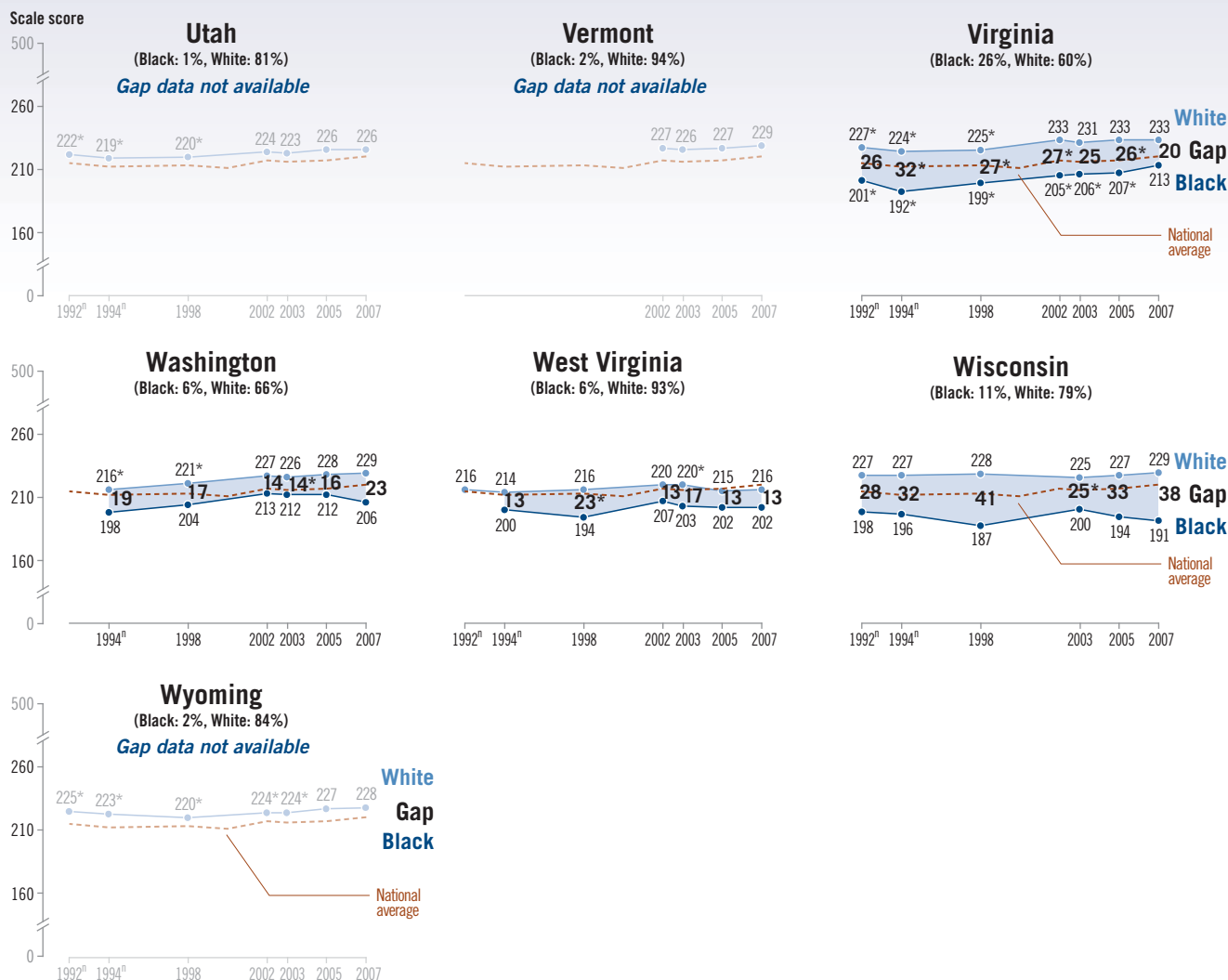
See notes at end of figure.

Figure 22. Gaps in average reading scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



See notes at end of figure.

Figure 22. Gaps in average reading scores between Black and White public school students at grade 4, by state: Various years, 1992–2007—Continued



ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

¹ National results for assessments prior to 2002 are based on the national sample, not on aggregated state samples.

² Black and White percentages are based on students tested in 2007.

³ Department of Defense Education Activity (overseas and domestic schools). Before 2005, DoDEA overseas and domestic schools were separate jurisdictions in NAEP. Pre-2005 data presented here were recalculated for comparability.

NOTE: Detail may not sum to totals due to rounding. Where data are not present, the jurisdiction did not participate or did not meet the minimum participation guidelines for reporting. State-level data were not collected in 2000. Comparative performance results may be affected by changes in exclusion rates for students with disabilities and English language learners in the NAEP samples.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Reading Assessments.

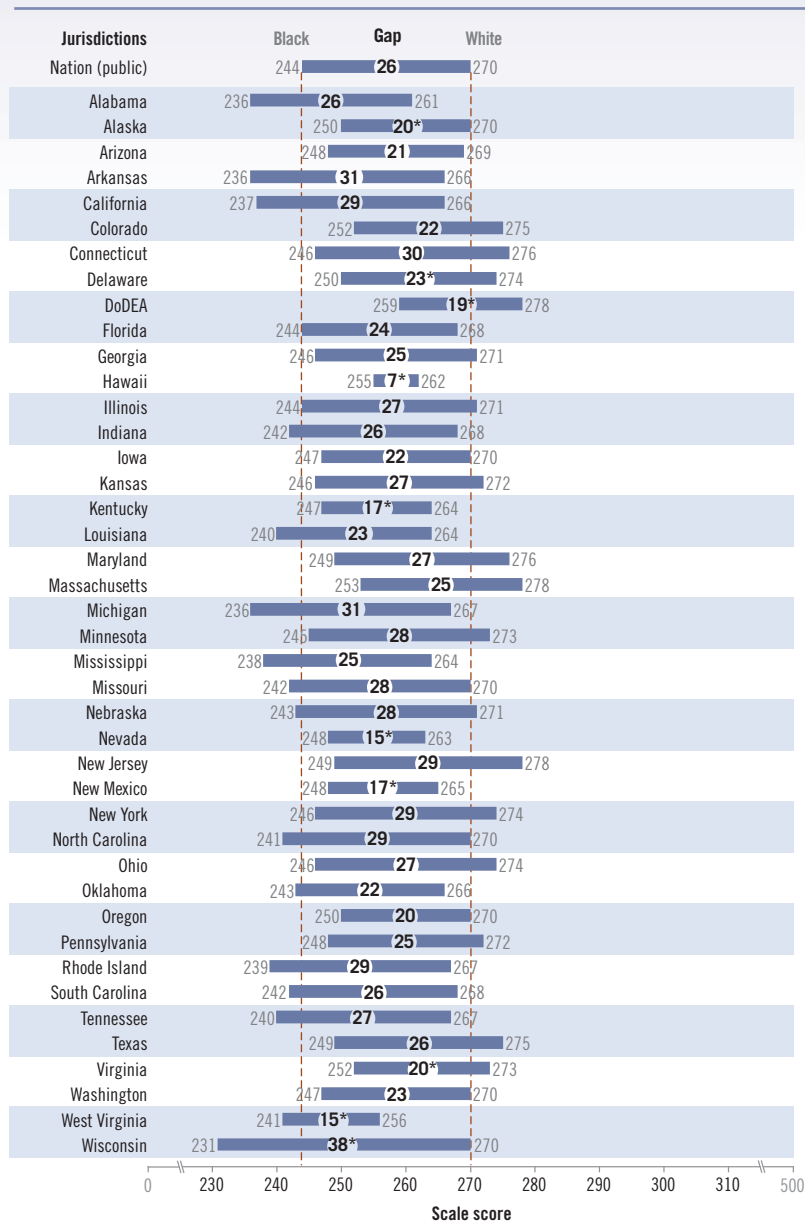


State and national reading achievement gaps at grade 8, 2007

Nine states had a Black-White gap that was smaller than the nation's 26-point gap in 2007 (Alaska, Delaware, DoDEA, Hawaii, Kentucky, Nevada, New Mexico, Virginia, and West Virginia) and one had a gap that was larger (Wisconsin). In 32 states, the gap was not significantly different from the nation's. State gaps that are either significantly larger or smaller than the national gap are indicated with asterisks (figure 23).

In Hawaii, the 7-point difference between the average scores for Black and White students was not statistically significant, and thus there was no Black-White gap for grade 8 reading in that state in 2007. In the other 41 states for which reliable data could be reported, the differences were statistically significant. The gaps ranged from 15 points in West Virginia and Nevada to 38 points in Wisconsin.

Figure 23. The Black-White achievement score gap in reading for public school students at grade 8, by state or jurisdiction: 2007



* Significantly different (p<.05) from the nation (public) when comparing one state to the nation at a time.

¹ Department of Defense Education Activity (overseas and domestic schools).

NOTE: States whose Black or White student population size was insufficient for comparison are omitted. Reporting standards not met for District of Columbia, Idaho, Maine, Montana, New Hampshire, North Dakota, South Dakota, Utah, Vermont, and Wyoming.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Reading Assessment.

Trends in state reading achievement gaps at grade 8, 1998–2007

Reading scores for the nation's public school students in the eighth grade were higher in 2007 than in 1992 for both Black and White students, but were not significantly different than in 1998 (figure 24, National results). Moreover, the national eighth-grade reading gap has not changed since either 1992 or 1998. NAEP first conducted eighth-grade reading assessments at the state level in 1998.

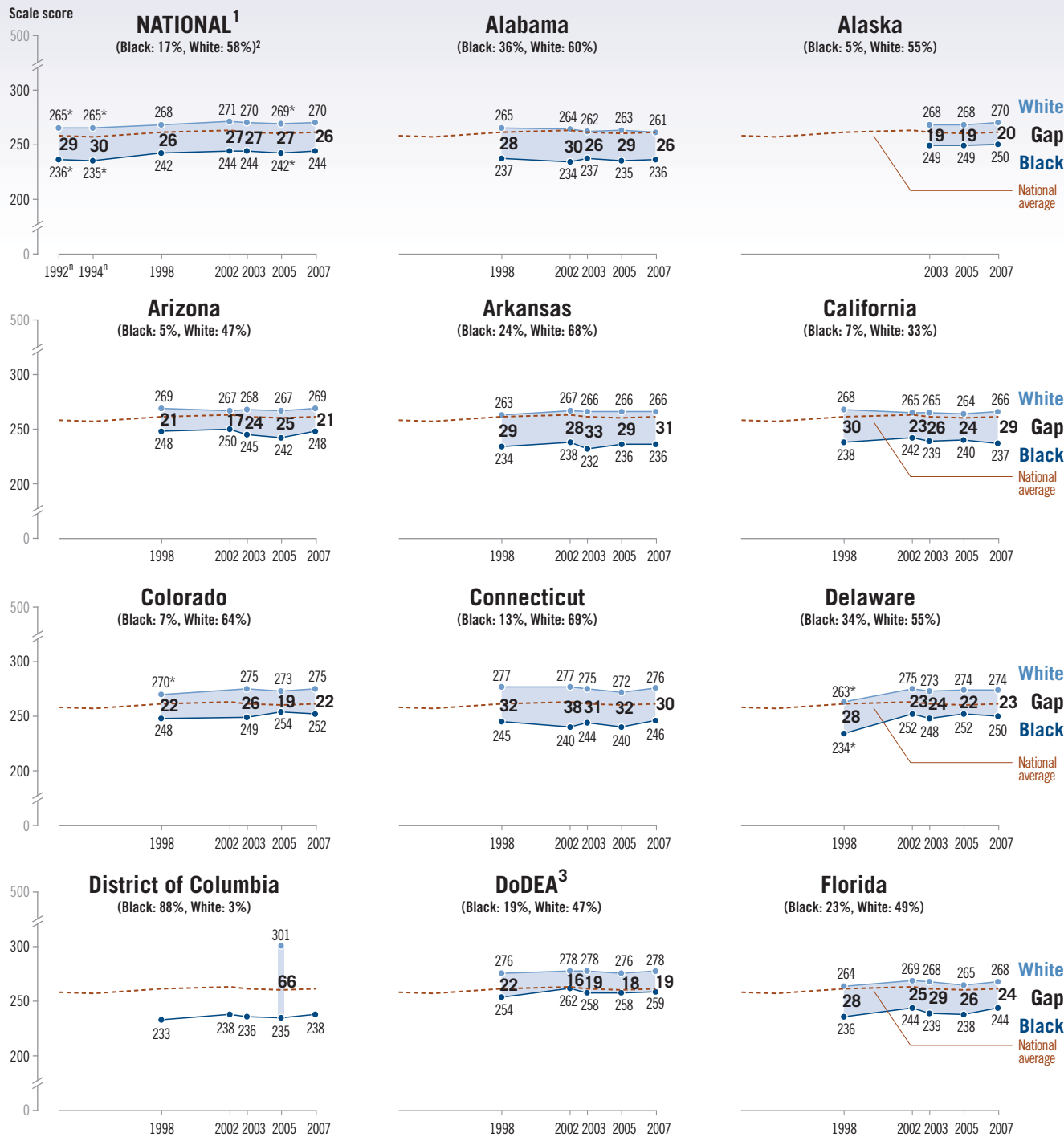
From 1998 to 2007, the Black-White score gap did not change for any state. In Delaware, scores for both Black and White eighth-graders were higher in 2007 than in 1998, but

there was no significant change in gap. During this period, five other states showed significant changes in average scores in one, but not both, of the two student groups. From 2005 to 2007, the gap did not change in any state.

Despite the fact that no statistically significant changes in state gaps were identified, the 7-point difference in Black and White students' scores in Hawaii for 2007 was itself not statistically significant, so that no Black-White score gap in grade 8 reading existed for that state.

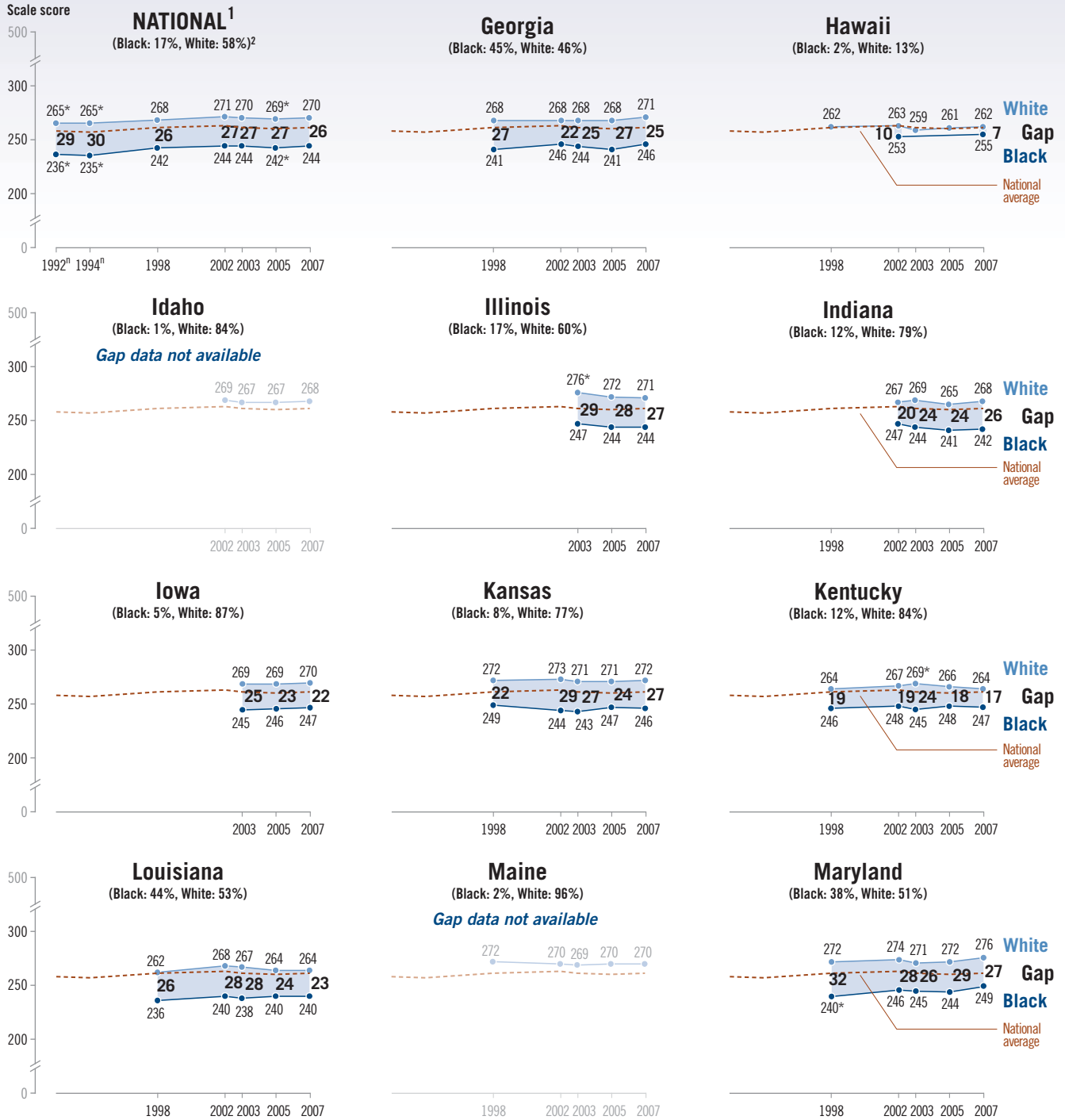


Figure 24. Gaps in average reading scores between Black and White public school students at grade 8, by state: Various years, 1998–2007



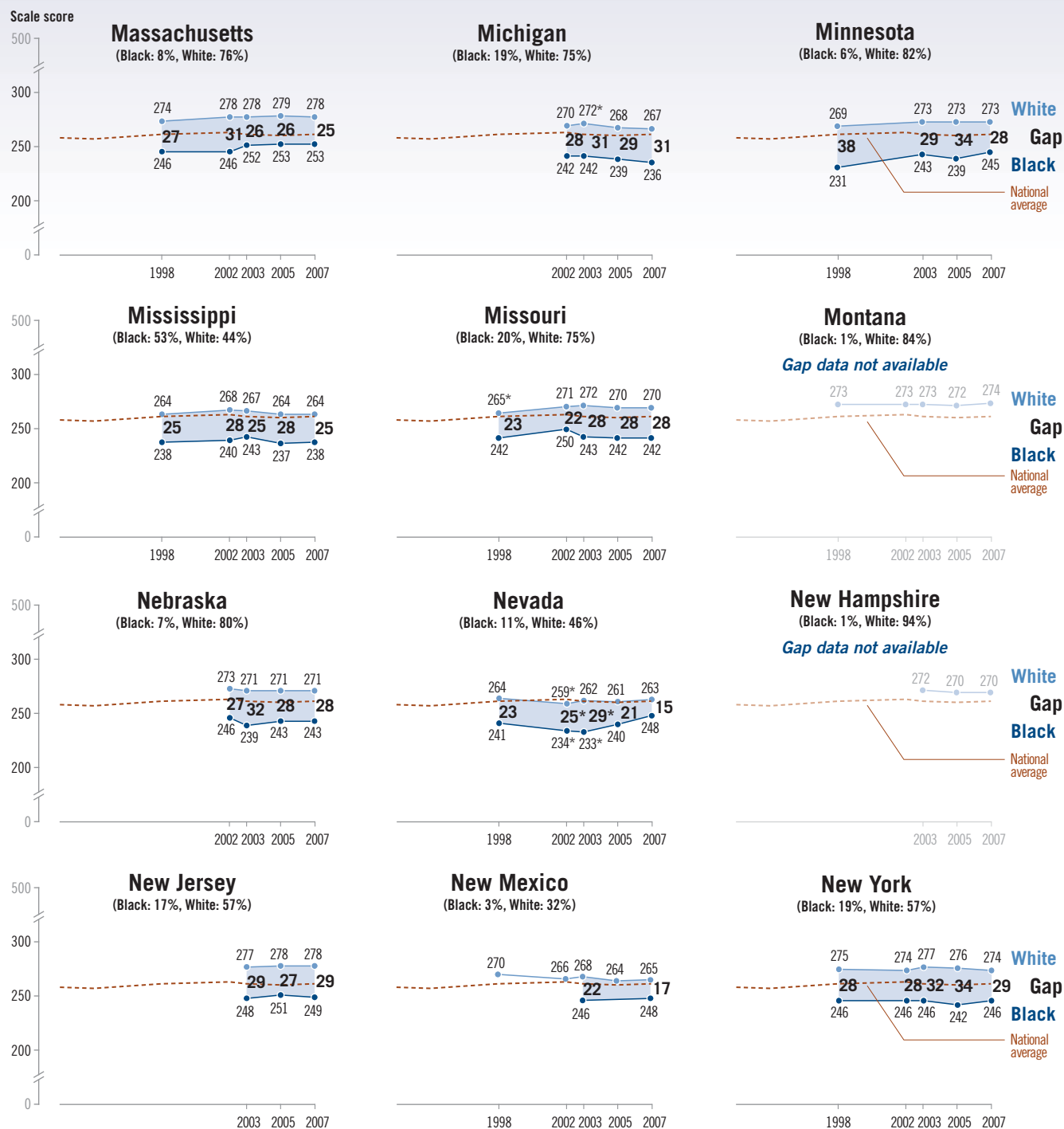
See notes at end of figure.

Figure 24. Gaps in average reading scores between Black and White public school students at grade 8, by state: Various years, 1998–2007—Continued



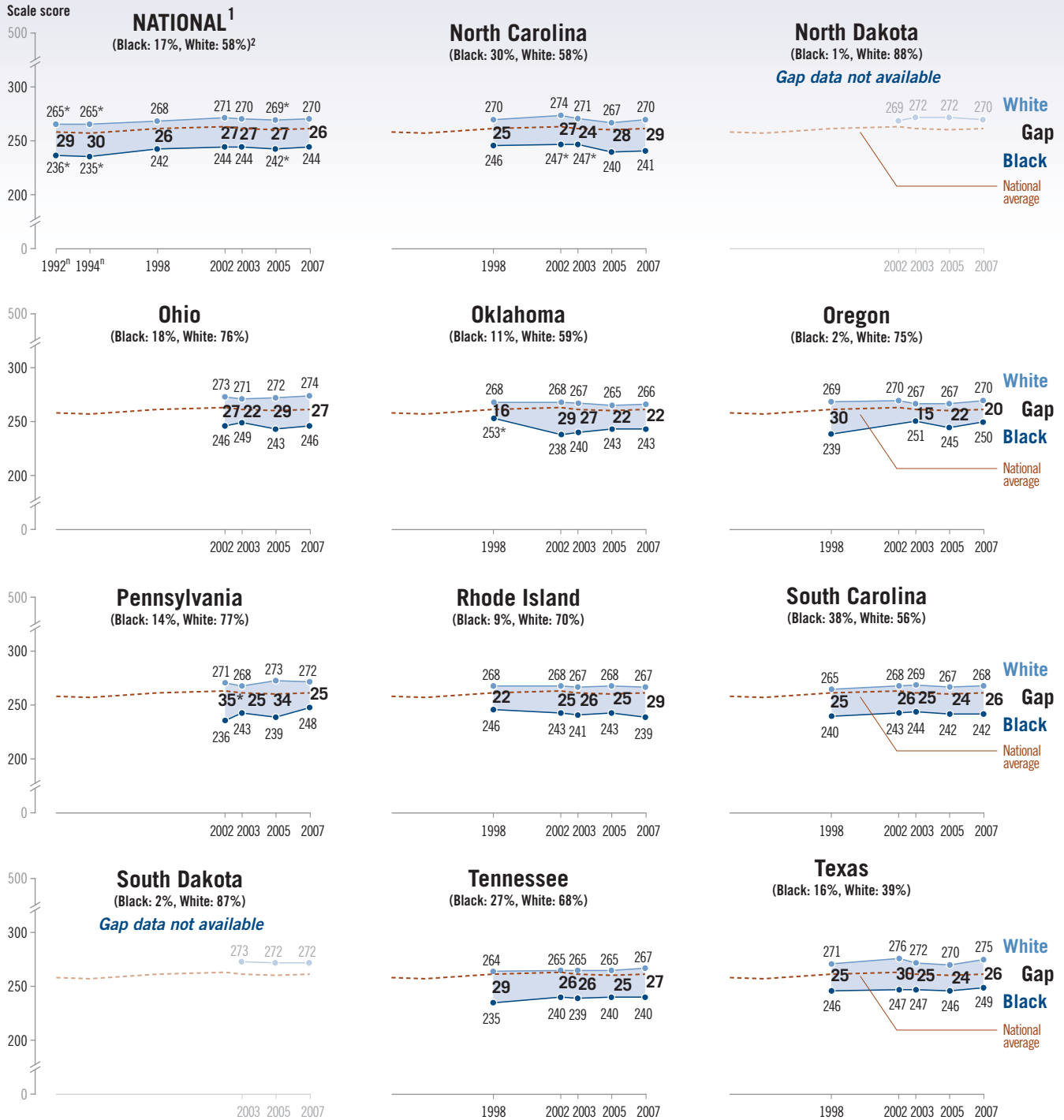
See notes at end of figure.

Figure 24. Gaps in average reading scores between Black and White public school students at grade 8, by state: Various years, 1998–2007—Continued



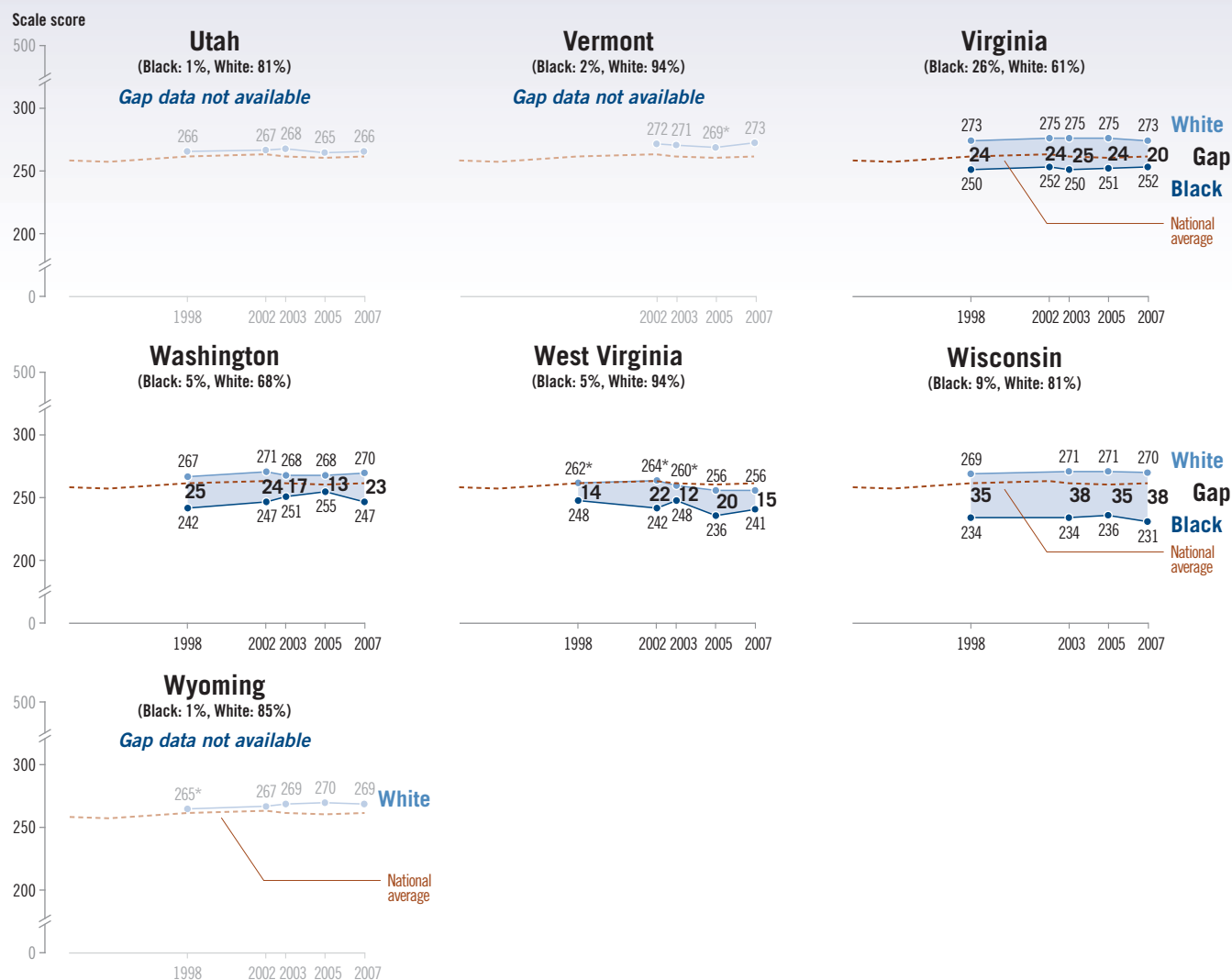
See notes at end of figure.

Figure 24. Gaps in average reading scores between Black and White public school students at grade 8, by state: Various years, 1998–2007—Continued



See notes at end of figure.

Figure 24. Gaps in average reading scores between Black and White public school students at grade 8, by state: Various years, 1998–2007—Continued



ⁿ Accommodations were not permitted for this assessment.

* Significantly different ($p < .05$) from 2007.

¹ National results for assessments prior to 2002 are based on the national sample, not on aggregated state samples.

² Black and White percentages are based on students tested in 2007.

³ Department of Defense Education Activity (overseas and domestic schools). Before 2005, DoDEA overseas and domestic schools were separate jurisdictions in NAEP. Pre-2005 data presented here were recalculated for comparability.

NOTE: Detail may not sum to totals due to rounding. Where data are not present, the jurisdiction did not participate or did not meet the minimum participation guidelines for reporting. State-level data were not collected in 1992, 1994, or 2000. Comparative performance results may be affected by changes in exclusion rates for students with disabilities and English language learners in the NAEP samples.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2007 Reading Assessments.

Appendix A: Technical Notes

This report presents data from two different assessment series, the NAEP long-term trend assessments and the main NAEP assessments. In most but not all cases, the two assessments used different procedures. Whenever a topic requires separate treatment of the two assessments, discussion of the long-term trend assessments, which present national results only, appears first, followed by the discussion of the main NAEP assessments, which present both national and state results. Discussion of main NAEP grade 12 assessments is omitted in this report because these assessments are conducted at the national level only.

Frameworks, development, administration, scoring, and analysis

Long-term trend

Overviews of these topics and more extensive information about other topics for the long-term trend assessments can be obtained from *NAEP 2004 Trends in Academic Progress*, available from the NAEP website <http://nces.ed.gov/nationsreportcard/ltr/>, which also provides links to earlier reports in the long-term trend series. (In 2004, the long-term assessments in reading and mathematics were conducted for two different “studies”: the “bridge study,” which was identical to previous long-term assessments, and the “modified study,” which will be used in future long-term assessments. The results for the 2004 assessment, reported in *NAEP 2004 Trends in Academic Progress* and in this report, were drawn from the bridge study.)

Main NAEP 2007 reading and mathematics assessments

For overviews of these topics, and for more extensive information about other topics for the 2007 main NAEP reading and mathematics assessments, consult the information available online at <http://nces.ed.gov/nationsreportcard/reading/> and <http://nces.ed.gov/nationsreportcard/mathematics/>

Sources of the data

Long-term trend

This report presents national data from the 1978, 1982, 1986, 1990, 1992, 1996, 1999, and 2004 long-term trend mathematics assessments and the 1980, 1984, 1990, 1992, 1994,

1996, 1999, and 2004 long-term trend reading assessments for Black and White public school students ages 9 and 13. Earlier long-term trend assessment results are available, but only for both public and private school students combined.

Main NAEP

This report presents national data from the 1990, 1992, 1996, 2000, 2003, 2005, and 2007 main NAEP mathematics assessments and the 1992, 1994, 1998, 2002, 2003, 2005, and 2007 main NAEP reading assessments for Black and White public school students in the fourth and eighth grades. In 2000, the reading assessment was also administered in the fourth grade (see tables B-1 and B-3 in appendix B).

This report presents state data from the 1992, 1996, 2000, 2003, 2005, and 2007 fourth-grade main NAEP mathematics assessments and from the 1990, 1992, 1996, 2000, 2003, 2005, and 2007 eighth-grade main NAEP mathematics assessments, for public school students only. The main NAEP reading assessment was administered at the state level to fourth-grade public school students in 1992, 1994, 1998, 2002, 2003, 2005, and 2007 and to eighth-grade public school students in 1998, 2002, 2003, 2005, and 2007.

Nationally in 2007, Black students constituted 17 percent of the public school fourth-grade population (based on data from the NAEP reading assessment) while White students constituted 56 percent. Results for the eighth-grade were similar: 17 percent and 58 percent, respectively. However, percentages vary widely between states. For example, Black students constituted a majority of the fourth-grade population in two states, the District of Columbia (84 percent in mathematics and 86 percent in reading) and Mississippi (52 percent in mathematics and 51 percent in reading). In contrast, Black students constituted only 2 percent of the fourth-grade public school population in states such as Wyoming and South Dakota. Eighth-grade data show a similar pattern. In some cases, the Black or White student population is so small that valid data cannot be obtained.

NAEP sampling procedures

Long-term trend

The populations sampled for the 2004 NAEP long-term trend assessment results presented in this report consisted

of 9- and 13-year-old students enrolled in public elementary and secondary schools nationwide. Eligibility for the age 9 and age 13 samples was based on calendar year: students in the age 9 sample were 9 years old on January 1, 2004, with birth months January 1994 through December 1994, and students in the age 13 sample were 13 years old on January 1, 2004, with birth months January 1990 through December 1990.

Consistent with past national long-term trend assessments, students were selected for participation based on a stratified three-stage sampling plan. In the first stage, geographic primary sampling units (PSUs) were defined and selected. In the second stage, schools were selected within PSUs. In the third stage, eligible students were selected within schools. Stratification occurred at both the school level and the PSU level. A full description of the sampling plan is beyond the scope of this appendix; for additional details regarding the design and structure of the 2004 trend assessment samples, the reader should refer to the technical documentation section of the NAEP website (<http://nces.ed.gov/nationsreportcard/ltt>).

The first-stage sampling units, PSUs, were drawn from a list—a sampling frame—developed using the metropolitan area designations of the U.S. Census Bureau. Each NAEP PSU in the frame was intended to encompass one county or contiguous multiple counties, generally not crossing state boundaries, and contained a minimum number of school-aged children—10,000 to 15,000, depending on the region of the country.

All PSUs containing more than 800,000 students (17 in all) were automatically included in the sample. Sixty additional PSUs were selected in a non-random manner, taking into account region of the country, status as either metropolitan or non-metropolitan, percentages of racial/ethnic groups, income levels, education levels in the population, and percentage of renters, with adjustments made to compensate for the non-random manner of selection.

In the second stage of sampling, schools were sampled from within the selected PSUs. Schools were selected with probability proportional to a measure of size based on the estimated number of age-eligible students in the school.

This in turn was estimated by applying population-level percentages of age-eligible students within each grade to

estimated grade enrollments for each grade, and aggregating to an age-eligible total for the school.

In the third stage of sampling, students were sampled from within schools. Sampled schools were asked to list all students with the appropriate birth dates for each specified age sample. All eligible students up to a pre-specified maximum (128 for both ages 9 and 13) were then selected for the assessment. If a school selected for the age 9 or age 13 samples had 128 or fewer students, all age-eligible students were selected into the sample for that school. Otherwise, a sample of 128 age-eligible students was taken.

The actual student and school sample sizes obtained in the NAEP long-term trend reading assessments, as well as the school and student participation rates, are presented in table A-1. Sample sizes and participation rates for the long-term trend mathematics assessments were similar. Although sampled schools that refused to participate were replaced, school participation rates were computed based on the schools originally selected for participation in the assessments. The student participation rates represent the percentage of students assessed of those invited to be assessed, including those assessed in follow-up sessions when necessary. Response rates for public school students ages 9 and 13 met NCES reporting standards for all assessments.

Table A-1. School and student participation rates, and target student population, Long-Term Trend Reading assessment, public school students only, by age: 2004

Participation and target population	Age 9	Age 13
School participation		
Weighted school percentage	88	85
Total number of schools that participated	250	230
Student participation		
Weighted student percentage	94	92
Total number of students who participated	3,800	4,000
Target population	3,700,000	3,690,000

NOTE: The numbers of schools are rounded to the nearest ten, the numbers of students are rounded to the nearest hundred, and the numbers for target populations are rounded to the nearest ten thousand. 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), 2004 Long-Term Trend Reading Assessments.

Main NAEP

The schools and students participating in NAEP assessments are chosen to be nationally representative. Samples of schools and students are selected from each state and from the District of Columbia and Department of Defense Education Activity (DoDEA) schools. The results from the assessed students are combined to provide accurate estimates of overall national performance and of the performance of individual states.

NCES has changed the main NAEP sampling methods over the years. From 1990 through 2000, the national sample was collected separately from the state samples. The 2002 national sample was the sum of all the state samples of the participating states, plus small samples from the few states that did not participate. In 2003, 2005, and 2007, all states participated and the national sample was the aggregate of the samples from all states, the District of Columbia, and the DoDEA schools. The main NAEP national samples in reading and mathematics since 2002 have been larger than in previous assessment years. Thus, smaller score differences between years or between types of student were found to be statistically significant than would have been detected in previous assessments.

From 1990 through 2001, NCES oversampled schools with high minority populations (Black and Hispanic) in the national sample. Beginning in 2002, this practice was discontinued because the state samples were large enough to ensure adequate coverage for these populations. Prior to 2002, NAEP results were weighted to compensate for the oversampling.

In 2003, 2005, and 2007, results were weighted to take into account the fact that states, and schools within states, represent different proportions of the overall national population. For example, since the number of students assessed in most states is roughly the same (to allow for stable state estimates and administrative efficiencies), the results for students in less populous states are assigned smaller weights than the results for students in more populous states. Sampling weights are also used to account for lower sampling rates for very small schools and are used to adjust for school and student nonresponse.

NAEP samples for reading and mathematics assessments administered from 1990 through 2007 are discussed in more detail below.

The NAEP 2007 mathematics and reading assessments were administered to fourth- and eighth-graders in all states. This report includes data for public school students for both the nation and all states. All 50 states, the District of Columbia, and the DoDEA schools met the minimum guidelines for reporting their results in 2007 for both assessments.

In order to obtain a representative sample for reporting national and state public school results in 2007, NCES sampled and assessed approximately 183,000 fourth-graders from 7,300 schools and 155,000 eighth-graders from 6,400 schools for the reading assessment and approximately 190,000 fourth-graders from 7,300 schools and 147,000 eighth-graders from 6,400 schools for the mathematics assessment.

Each selected school that participated in the assessment and each student assessed represent a portion of the population. The schools were selected out of approximately 51,000 fourth-grade and 27,000 eighth-grade public schools. The students selected from these schools represented the total population of approximately 3.4 million fourth-grade and 3.6 million eighth-grade public school students. These totals include the public schools in the 50 states and the District of Columbia.

Schools in the DoDEA school system are classified as “nonpublic” by NCES and their results are not included in the determination of NAEP national public average scale scores. These schools are not “private” because they are operated by the federal government and they are not “public” because only children of U.S. military personnel can attend them. For comparison purposes, the system is treated as a state and results are compared with the scores of the 50 states and the District of Columbia.

Table A-2 provides a summary of the 2007 national and state school and student participation rates for the reading grade 8 assessment sample. Rates for reading grade 4 and mathematics grades 4 and 8 in 2007 were similar, as

were the rates for the 2003 and 2005 assessments. Readers who want more detail should consult the 2007, 2005 and 2003 report cards, available online at <http://nces.ed.gov/pubsearch/getpubcats.asp?sid=031>.

Participation rates in table A-2 are presented for public schools and public school students in grade 8 reading. The school participation rate is a school-centered, weighted percentage of schools participating in the assessment. This rate is based only on the schools that were initially selected for the assessment. The numerator of this rate is the estimated number of schools represented by the initially selected schools that participated in the assessment. The denominator is the estimated number of schools represented by the initially selected schools that had eligible students enrolled.

Also presented in table A-2 are weighted student participation rates. The numerator of this rate is the estimated number of students who are represented by the students assessed (in either an initial session or a makeup session). The denominator of this rate is the estimated number of students represented by the eligible sampled students in participating schools.

The term “eligible students” used in the two preceding paragraphs refers to students who can meaningfully participate in NAEP. Students excluded from NAEP assessments on the grounds that they cannot meaningfully participate—whether students with disabilities or English language learners—are not part of the population of interest. Initially selected schools that had no eligible students enrolled are excluded from the denominator of the school participation rate because they contained no students who were part of the population of interest. For similar reasons, the denominator of the weighted student participation rate consists only of eligible sampled students.

The fourth column gives the number of public school students who were assessed in each of the jurisdictions. The final column of table A-2 gives the target populations for each jurisdiction, that is, the eighth-grade population for that jurisdiction.

The national target population per grade for all main NAEP assessments 1990–2007 ranged from about 3.25 million

to about 3.75 million. In the 1990–1996 assessments, the number of schools sampled per assessment and grade for the national sample ranged from approximately 120 to 230, while the number of students assessed ranged from approximately 5,200 to 9,900. In the 1998–2000 assessments, the number of schools sampled per assessment and grade ranged from approximately 330 to 390, while the number of students assessed ranged from approximately 6,100 to 9,000.

The state target populations for all main NAEP assessments 1990–2007 ranged from approximately 5,000 in the District of Columbia and 9,000 in sparsely populated states like Wyoming and Alaska to approximately 450,000 in California, followed by approximately 325,000 in Texas.

In the 1990–2000 state assessments, the number of schools sampled per assessment and grade ranged from approximately 30 to 150, while the number of students assessed ranged from approximately 1,000 to 5,900. In the 2003–2007 state assessments, the number of schools sampled per assessment and grade ranged from approximately 40 to 250, while the number of students assessed ranged from approximately 1,700 to 10,700.

In earlier NAEP assessments, NCES would select substitute schools that would be used to augment the original sample if a large number of schools from the sample failed to participate. School and student participation rates were given both before and after substitution. Because the No Child Left Behind Act requires states to participate in the main NAEP reading and mathematics assessments at the fourth and eighth grades in order to qualify for full Title I education funding, participation rates are very high and NCES no longer selects substitute schools for these assessments.

In order to ensure unbiased samples, NCES and the National Assessment Governing Board, which establishes policy for NAEP, set minimums for the school participation rate before substitution of replacement schools for any sample. From 1990 through 2002, the standard for the state assessments required that the weighted school participation rate before substitution of replacement schools

Table A-2. School and student participation rates, and target student population, grade 8 reading assessment, public school students only, by state or jurisdiction: 2007

Jurisdiction	School participation		Student participation		Target population
	Weighted school percentage	Total number of schools that participated	Weighted student percentage	Total number of students who participated	
Nation (public)	100	6,410	92	154,700	3,558,000
Alabama	100	120	93	2,800	56,000
Alaska	99	110	91	2,600	9,000
Arizona	100	130	90	2,800	73,000
Arkansas	100	120	93	2,500	34,000
California	100	310	92	8,600	477,000
Colorado	98	120	92	2,800	57,000
Connecticut	97	100	92	2,700	42,000
Delaware	100	50	93	2,800	10,000
District of Columbia	100	50	88	1,800	5,000
DoDEA ¹	98	60	94	1,700	5,000
Florida	100	160	91	4,100	193,000
Georgia	100	120	93	3,500	120,000
Hawaii	100	70	91	2,800	13,000
Idaho	99	110	93	2,900	20,000
Illinois	100	200	93	4,000	150,000
Indiana	100	110	92	2,700	80,000
Iowa	100	130	93	2,800	36,000
Kansas	100	150	94	2,800	34,000
Kentucky	100	110	93	2,600	46,000
Louisiana	100	110	92	2,400	47,000
Maine	98	130	93	2,700	15,000
Maryland	100	110	90	2,700	64,000
Massachusetts	100	140	93	3,600	70,000
Michigan	100	120	91	2,600	119,000
Minnesota	99	140	92	3,000	62,000
Mississippi	100	110	93	2,700	36,000
Missouri	100	130	92	2,900	70,000
Montana	98	170	92	2,600	11,000
Nebraska	100	120	94	2,700	21,000
Nevada	100	70	88	2,600	28,000
New Hampshire	98	90	92	2,900	16,000
New Jersey	97	110	92	2,800	104,000
New Mexico	100	110	89	2,600	25,000
New York	100	160	90	3,800	206,000
North Carolina	100	150	91	4,300	104,000
North Dakota	98	190	95	2,200	8,000
Ohio	100	190	92	3,500	135,000
Oklahoma	100	150	92	2,600	42,000
Oregon	100	110	92	2,700	39,000
Pennsylvania	100	110	92	2,800	140,000
Rhode Island	100	60	92	2,800	12,000
South Carolina	100	110	94	2,700	52,000
South Dakota	99	140	95	2,800	10,000
Tennessee	100	120	92	2,800	74,000
Texas	100	220	92	7,100	294,000
Utah	100	100	91	2,800	36,000
Vermont	100	120	93	2,000	7,000
Virginia	100	110	93	2,800	91,000
Washington	100	130	91	3,000	78,000
West Virginia	100	120	92	2,900	21,000
Wisconsin	98	130	92	2,700	62,000
Wyoming	100	80	92	2,000	7,000

¹ Department of Defense Education Activity (overseas and domestic schools).

NOTE: The numbers of schools are rounded to the nearest ten, the numbers of students are rounded to the nearest hundred, and the target population is rounded to the nearest thousand. 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), 2007 Reading Assessment.

be 70 percent or higher. Beginning in 2003, the standard was raised to 85 percent. All data presented in this report are based on samples meeting the standards in effect at the time of the assessment.

Since 1990, the national weighted public school participation rate before substitution for the grade 4 and 8 reading and mathematics assessments has ranged from 76 percent to 100 percent. Prior to 2003, a few states did not meet the 70 percent standard. From 1990 through 2002, the weighted public school participation rate before substitution for states whose results are reported here ranged from 70 percent to 100 percent.

For more information on all the NAEP assessments referenced in this report, consult the individual reports devoted to them, available from the NCES website at <http://nces.ed.gov/pubsearch/getpubcats.asp?sid=031>.

Understanding NAEP reporting groups

NAEP results are provided for groups of students defined by shared characteristics—race/ethnicity, eligibility for free/reduced-price school lunch, and gender, for example.

Based on participation rate criteria, results are reported for groups only when sufficient numbers of students and adequate school representation are present. The minimum requirement is a total of at least 62 students in a particular group, assessed in at least five different locations. However, the data for all students, regardless of whether their group was reported separately, were included in computing over-all national results. Definitions of the student groups discussed in this report follow.

Race/ethnicity

Long-term trend

In long-term trend NAEP, data about student race/ethnicity is based on the assessment administrator's observation. Self-reported race/ethnicity data has been collected since 1984, and school records-based race/ethnicity data has been collected starting in 2004, but all long-term trend results are reported based on observed race/ethnicity.

Main NAEP

In all main NAEP assessments, data about student race/ethnicity are collected from two sources: school records and student self-reports. In this report, the race/ethnicity variable has been based on the race reported by the school for all assessment years. In the rare cases when school-recorded information is missing, student-reported data are used to determine race/ethnicity.

Schools sampled for NAEP are asked to provide lists of all students in the target grade(s) along with basic demographic information, including race/ethnicity. Students are categorized into one of five mutually exclusive racial/ethnic categories plus “other.” Administration schedules—also referred to as student rosters—are created that include the list of sampled students along with their basic demographic information. These data are checked and updated during data collection. This race/ethnicity information is available for all sampled students: those that participated and those that were absent or excluded.

All students who take a NAEP assessment complete a section of general student background questions, including questions about their race/ethnicity. Separate questions are asked about students' Hispanic ethnic background and about students' race. This race/ethnicity information is available just for students who participated in the assessment and not for those who were absent or excluded. See <http://nces.ed.gov/nationsreportcard/bgquest.asp> for more information.

The mutually exclusive racial/ethnic categories are White (non-Hispanic), Black (non-Hispanic), Hispanic, Asian/Pacific Islander, American Indian (including Alaska Native), and Unclassified. Unclassified students are those whose school-reported race was “other,” or “unavailable,” or was missing, or who self-reported more than one race category (i.e., “multi-racial”) or none. Hispanic students may be of any race. Only results for White (non-Hispanic) and Black (non-Hispanic) students are contained in this report. Information based on student self-reported race/ethnicity is available on the NAEP Data Explorer (<http://nces.ed.gov/nationsreportcard/nde>).

Eligibility for free/reduced-price school lunch

Long-term trend

The long-term trend assessments do not report results based on school lunch eligibility.

Main NAEP

As part of the Department of Agriculture's National School Lunch Program, schools can receive cash subsidies and donated commodities in return for offering free or reduced-price lunches to eligible children. Based on available school records, students were classified as currently eligible for either free lunch or reduced-price lunch, or not eligible. Eligibility for the program is determined by a student's family income in relation to the federally established poverty level. Free lunch qualification is set at 130 percent of the poverty level or below, and reduced-price lunch qualification is set at between 130 and 185 percent of the poverty level. (For the period July 1, 2006, through June 30, 2007, for a family of four, 130 percent of the poverty level was \$26,000, and 185 percent was \$37,000. See <http://www.fns.usda.gov/cnd/lunch> for more information.) The classification applies only to the school year when the assessment was administered and is not based on eligibility in previous years. If school records were not available, the student was classified as "Information not available." If the school did not participate in the program, all students in that school were classified as "Information not available." As a result of improvements in the quality of the data on students' eligibility for NSLP, the percentage of students for whom information was not available has decreased in comparison to the percentages reported prior to the 2003 assessment. Therefore, trend comparisons are only made back to 2003 in this report.

Gender

Both long-term trend and NAEP assessments identify students as male or female based on school records.

Inclusion and exclusion

Long-term trend

Some students selected for participation in the NAEP long-term trend assessments were identified as English language learners (ELL) or students with disabilities (SD). In all previous long-term trend assessments, if it was decided that a student classified as SD or ELL could not meaningfully participate in the NAEP assessment for which he or she was selected, the student was, according to NAEP guidelines, excluded from the assessment.

For each student selected to participate in NAEP who was identified as either SD or ELL, a member of the school staff most knowledgeable about the student completed an SD/ELL questionnaire. Students with disabilities were excluded from the assessment if an individualized education program (IEP) team or equivalent group determined that the student could not participate in assessments such as NAEP; if the student's cognitive functioning was so severely impaired that the student could not participate; or if the student's IEP required that the student be tested with an accommodation or adaptation not permitted or available in NAEP, and the student could not demonstrate his/her knowledge of the assessment subject area without that accommodation or adaptation. A student who was identified as ELL and who was a native speaker of a language other than English was excluded if the student had received instruction in the assessment's subject area (e.g., reading or mathematics) primarily in English for less than three school years, including the current year, or if the student could not demonstrate his or her knowledge of reading or mathematics in English without an accommodation or adaptation.

Prior to 2004, NAEP long-term trend assessments did not allow accommodations for SD or ELL students. In that year, two versions of the long-term trend assessment were given, the "bridge" (unmodified) version, which did not allow accommodations, and the "modified" version, which

did. In 2004, results were only reported for the bridge assessment and all results from the 2004 Long-Trend Assessment appearing in this report are drawn from the bridge assessment. However, table A-3 presents exclusion rates for both versions of the 2004 assessment in order to give all the available information on the 2004 exclusion rates for Black and White students.

In the 2004 bridge assessment, and in all prior administrations of the long-term trend assessment, student race/ethnicity was determined by NCES contractor staff administering the assessment in the individual classrooms. These staff never met the excluded students, so no records of the race/ethnicity of excluded students were kept.

In contrast, the 2004 modified assessment determined student race/ethnicity by using school records, which did provide information on the race/ethnicity of excluded students. Exclusion data from the 2004 modified assessment are provided here to provide information on 2004 exclusion rates for Black and White students, even though this report does not include student achievement data drawn from the modified assessment.

Main NAEP

The NAEP program has always endeavored to assess all students selected as a part of its sampling process. In all NAEP schools, accommodations will be provided as necessary for students with disabilities (SD) and/or English language learners (ELL) or limited English proficient (LEP) students. (ELL is the term used since the NAEP 2005 reports; LEP was used before 2005.) The accommodations are available to students whose Individualized Education Program (IEP) specifically requires them. Because some ELL students do not have an IEP, decisions about accommodations for these students are typically made by knowledgeable school staff.

The NAEP program has established procedures to include as many SD and ELL students as possible in the assessments. School staff make the decisions about whether to include such a student in a NAEP assessment, and which testing accommodations, if any, they should receive. The NAEP program furnishes tools to assist school personnel in making those decisions.

A sampling procedure is used to select students at each grade being tested. Students are selected on a random basis, without regard to SD or ELL status. Once the students are selected, the schools identify which have SD or ELL status. School staff who are familiar with these students are asked a series of questions to help them decide whether each student should participate in the assessment and whether the student needs accommodations.

Inclusion in NAEP of an SD or ELL student is encouraged if that student (a) participated in the regular state academic assessment in the subject being tested, and (b) if that student can participate in NAEP with the accommodations NAEP allows. Even if the student did not participate in the regular state assessment, or if he/she needs accommodations NAEP does not allow, school staff are asked whether that student could participate in NAEP with the allowable accommodations.

History of NAEP Inclusion Policy Although NAEP has always endeavored to assess as high a proportion of

Table A-3. National Long-Term Trend mathematics and reading exclusion rates as percentages of the total sample, by age, type of assessment and race/ethnicity: 2004

	Age 9		Age 13	
	Bridge	Modified	Bridge	Modified
Mathematics				
Total	8	3	9	3
White	†	2	†	3
Black	†	4	†	4
Reading				
Total	9	6	9	5
White	†	4	†	5
Black	†	4	†	6

† Not applicable.

NOTE: The 2004 bridge assessment, and all previous administrations of the long-term trend assessment, did not obtain information on the race/ethnicity of excluded students.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2004 Long-Term Trend Mathematics and Reading Assessments.

sampled students as is possible, prior to 1996 NAEP did not allow accommodations for SD or ELL students. This resulted in exclusion of some students who could not meaningfully participate in the assessment without accommodations.

The passage of the Individuals with Disabilities Education Act (IDEA), as amended in 1997, led states and districts to identify increasing numbers of students as requiring accommodations in assessments in order to fairly and accurately show their abilities. It was important for NAEP to be as consistent as possible with testing practices in most states and districts while maintaining the ability to compare more recent NAEP results to those from 1990, 1992, and 1994, when accommodations were not allowed. (Accommodations were not allowed in NAEP state assessments until 1996.)

Before the 2005 assessment (when the selection process was detailed in a series of questions), guidelines were specified by NAEP. A student identified on the Administration Schedule as having a disability (SD), that is, a student with an Individualized Education Program (IEP) or equivalent classification, should be included in the NAEP assessment unless:

- The IEP team or equivalent group had determined that the student could not participate in assessments such as NAEP, or
- The student’s cognitive functioning was so severely impaired that he or she could not participate, or
- The student’s IEP required that the student be tested with an accommodation that NAEP did not permit, and the student could not demonstrate his or her knowledge of the subject without that accommodation.

A student who was identified as LEP or ELL and who was a native speaker of a language other than English should be included in the NAEP assessment unless:

- The student had received reading or mathematics instruction primarily in English for less than 3 school years including the current year, and
- The student could not demonstrate his or her knowledge of the subject in English even with an accommodation permitted by NAEP.

The phrase “less than 3 school years including the current year” meant 0, 1, or 2 school years. Therefore, the guidelines below were used:

- Include without any accommodation all LEP or ELL students who had received instruction in the subject primarily in English for 3 years or more and those who were in their third year;
- Include without any accommodation all other such students who could demonstrate their knowledge of the subject without an accommodation;
- Include and provide accommodations permitted by NAEP to other such students who can demonstrate their knowledge of the subject only with those accommodations; and
- Exclude LEP or ELL students only if they could not demonstrate their knowledge of the subject even with an accommodation permitted by NAEP.

The percentages of students excluded from NAEP may vary from one state to another, as well as across years. National exclusion rates for Black and White SD and/or ELL students in 2007 may be found in table A-4. The “total” rates include all students, not just those who are Black or White. For information on state exclusion rates, see table A-5. For more information on Main NAEP inclusion and exclusion, go to <http://nces.ed.gov/nationsreportcard/about/inclusion.asp>

Table A-4. National mathematics and reading exclusion rates as percentages of the total sample, public schools only, by grade and race/ethnicity: 2007

Grade and race/ethnicity	Mathematics	Reading
Grade 4		
Total	3	6
White	2	4
Black	4	7
Grade 8		
Total	4	6
White	4	4
Black	6	7

NOTE: “Total” exclusion percentages are for all public school students, not just Black and White.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Mathematics and Reading Assessments.

Table A-5. Mathematics and reading exclusion rates as percentages of the total sample, public schools only, by grade, race/ethnicity and jurisdiction: 2007

Jurisdiction	Percentage of students with a disability and/or English language learner, excluded in 2007							
	Grade 4 Mathematics		Grade 8 Mathematics		Grade 4 Reading		Grade 8 Reading	
	White	Black	White	Black	White	Black	White	Black
Nation (public)	2	4	3	6	4	7	4	7
Alabama	1	2	2	4	3	3	3	4
Alaska	1	2	4	6	3	3	2	2
Arizona	2	4	2	4	3	4	3	6
Arkansas	2	5	2	3	5	8	4	7
California	1	3	1	3	2	5	2	4
Colorado	2	4	1	2	2	7	2	4
Connecticut	1	1	1	1	2	3	2	4
Delaware	4	6	5	8	9	13	5	8
District of Columbia	2	5	‡	10	6	12	‡	13
DoDEA	1	1	2	1	4	5	2	3
Florida	1	4	2	3	4	6	3	5
Georgia	2	3	4	5	7	8	6	7
Hawaii	1	1	1	‡	3	1	2	0
Idaho	1	‡	1	‡	3	‡	3	‡
Illinois	3	5	4	10	5	6	4	6
Indiana	2	4	5	9	3	5	4	8
Iowa	1	2	2	6	3	13	4	15
Kansas	2	6	4	4	4	8	4	7
Kentucky	3	3	7	5	7	10	8	7
Louisiana	1	3	2	4	3	5	2	4
Maine	3	4	5	‡	6	‡	5	‡
Maryland	2	5	6	9	5	10	4	10
Massachusetts	4	6	8	12	5	9	6	10
Michigan	3	4	4	7	4	6	5	9
Minnesota	2	3	2	3	3	9	3	10
Mississippi	1	1	1	3	2	3	2	4
Missouri	3	4	5	6	4	4	3	4
Montana	2	‡	3	‡	4	‡	4	‡
Nebraska	2	5	2	3	4	8	3	2
Nevada	2	6	3	4	5	7	3	9
New Hampshire	2	8	3	13	4	13	3	‡
New Jersey	1	4	2	6	5	8	4	10
New Mexico	2	4	2	3	6	9	4	11
New York	2	2	4	4	5	5	5	4
North Carolina	1	3	1	3	1	3	2	3
North Dakota	4	‡	5	‡	7	34	8	‡
Ohio	4	8	6	12	7	12	8	10
Oklahoma	5	4	8	11	6	7	6	10
Oregon	2	5	3	5	4	7	2	5
Pennsylvania	2	3	3	4	4	6	4	8
Rhode Island	1	4	2	5	3	4	3	6
South Carolina	1	2	4	7	4	4	5	7
South Dakota	1	1	2	‡	5	17	5	‡
Tennessee	5	7	6	7	9	12	7	8
Texas	5	6	4	7	6	8	4	8
Utah	2	‡	2	‡	5	‡	4	‡
Vermont	2	‡	4	‡	6	29	5	‡
Virginia	4	6	4	9	7	9	6	9
Washington	2	3	2	7	4	6	3	10
West Virginia	1	1	2	5	2	1	2	4
Wisconsin	2	4	3	9	3	8	5	12
Wyoming	1	‡	2	‡	3	‡	3	‡

‡ Reporting standards not met.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Mathematics and Reading Assessments.

Accommodations

Long-term trend

The long-term trend results presented in this report are drawn from assessments that did not permit accommodations for students with disabilities (SD) and English language learners (ELL). Future long-term trend assessments will allow such accommodations.

Main NAEP

From 1990 through 1994 for the nation—and through 1996 for the states—main NAEP assessments did not allow accommodations for either SD or ELL students. Since then, accommodations have been permitted for those SD and ELL students who need accommodations in order to participate, unless the accommodation would change the nature of what is being tested.

To accomplish this goal, students who receive accommodations in their state's assessments are offered the same accommodations on NAEP, except where an accommodation would change the nature of what is being tested. For example, passages and questions in the reading test are not permitted to be read aloud to the student, because that accommodation would make it a test of listening instead of a test of reading. Similarly, reading passages and questions cannot be presented in a language other than English.

It should be noted that students assessed with accommodations typically received some combination of accommodations. For example, students assessed in small groups (as compared with standard NAEP sessions of about 30 students) usually received extended time. In one-on-one administrations, students often received assistance in recording answers (e.g., use of a scribe or computer) and were afforded extra time.

The most common accommodations are small-group administration, extended time, one-on-one administration, the use of a scribe or computer, and the use of a bilingual book (mathematics only). See <http://nces.ed.gov/nationsreportcard/tdw/instruments/accomm.asp> for more details on NAEP accommodations. For state accommodation rates for SD and ELL students in 2007 see

the Technical Notes sections of *The Nation's Report Card: Mathematics 2007* at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007494> and *The Nation's Report Card: Reading 2007* at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007496>.

Drawing inferences from the results

The reported statistics for both long-term trend and main NAEP are estimates and are therefore subject to a measure of uncertainty. There are two sources of such uncertainty. First, NAEP uses a sample of students rather than testing all students. Second, all assessments have some amount of uncertainty related to the fact that they cannot ask all questions that might be asked in a content area. The magnitude of this uncertainty is reflected in the standard error of each of the estimates. When the percentages or average scale scores of certain groups are compared, the estimated standard error should be taken into account. Therefore, the comparisons are based on statistical tests that consider the estimated standard errors of the statistics being compared and the magnitude of the difference between the averages or percentages.

Standard errors for the NAEP scores and percentages presented in this report for both assessments are available on the NAEP website (<http://nces.ed.gov/nationsreportcard/naepdata>).

The differences between statistics—such as comparisons of two groups of students' average scale scores—that are discussed in this report are determined by using standard errors. Comparisons are based on statistical tests that consider both the size of the differences and the standard errors of the two statistics being compared. Estimates based on smaller groups are likely to have relatively large standard errors. As a consequence, a numerical difference that seems large may not be statistically significant.

Furthermore, differences of the same magnitude may or may not be statistically significant, depending upon the size of the standard errors of the statistics. For example, a 3-point change in the gap between Black and White fourth-graders nationwide may be significant, while a

3-point change in the gap between Black and White fourth-graders in Kansas may not be. The differences described in this report have been determined to be statistically significant at the .05 level with appropriate adjustments for part-to-whole and multiple comparisons.¹

In the tables and figures of this report, the symbol (*) is used to indicate that a score or percentage is significantly different from another. In addition, any difference between scores or percentages that is identified as higher, lower, larger, smaller, narrower, or wider in this report, including within-group differences not marked in tables and figures, meets the requirements for statistical significance.

Weighting and variance estimation

In both long-term trend and main NAEP a complex sample design was used to select the students who were assessed. The properties of a sample selected through such a design could be very different from those of a simple random sample, in which every student in the target population has an equal chance of selection and in which the observations from different sampled students can be considered to be statistically independent of one another. Therefore, the properties of the sample for the data collection design were taken into account during the analysis of the assessment data.

One way that the properties of the sample design were addressed was by using sampling weights to account for the fact that the probabilities of selection were not identical for all students. All population and subpopulation characteristics based on the assessment data were estimated using sampling weights. These weights included adjustments for school and student nonresponse.

Not only must appropriate estimates of population characteristics be derived, but appropriate measures of the degree of uncertainty must be obtained for those statistics. Two components of uncertainty are accounted for in the variability of statistics based on student ability: (1) the uncertainty due to sampling only a relatively small number of students,

and (2) the uncertainty due to sampling only a relatively small number of cognitive questions. The first component accounts for the variability associated with the estimated percentages of students who had certain background characteristics or who answered a certain cognitive question correctly.

Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate. NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any student information that can be observed without error. However, because each student typically responds to only a few questions within a content area, the scale score for any single student would be imprecise. In this case, NAEP's marginal estimation methodology can be used to describe the performance of groups and subgroups of students. The estimate of the variance of the students' posterior scale score distributions (which reflect the imprecision due to lack of measurement accuracy) is computed. This component of variability is then included in the standard errors of NAEP scale scores.²

Analyzing group differences in averages and percentages

In both long-term trend and main NAEP, statistical tests determine whether, based on the data from the groups in the sample, there is strong enough evidence to conclude that the averages or percentages are actually different for those groups in the population. If the evidence is strong (i.e., the difference is statistically significant), the report describes the group averages or percentages as being different (e.g., one group performed higher or lower than another group), regardless of whether the sample averages or percentages appear to be approximately the same. The reader is cautioned to rely on the results of the statistical tests rather than on the apparent magnitude of the difference between sample averages or percentages when

¹ Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society, Series B*, no. 1, 289–300.

² For further detail, see Johnson, E.G., and Rust, K.F. (1992). Population Inferences and Variance Estimation for NAEP Data. *Journal of Educational Statistics*, (17)2, 175–190.

determining whether the sample differences are likely to represent actual differences among the groups in the population.

To determine whether a real difference exists between the average scale scores (or percentages of a certain attribute) for two groups in the population, one needs to obtain an estimate of the degree of uncertainty associated with the difference between the averages (or percentages) of these groups for the sample. This estimate of the degree of uncertainty, called the “standard error of the difference” between the groups, is obtained by taking the square of each group’s standard error, summing the squared standard errors, and taking the square root of that sum.

$$SE_{A-B} = \sqrt{(SE_A^2 + SE_B^2)}$$

The standard error of the difference can be used, just like the standard error for an individual group average or percentage, to help determine whether differences among groups in the population are real. The difference between the averages or percentages of the two groups plus or minus 1.96 standard errors of the difference represents an approximately 95 percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between the groups in the population. If the interval does not contain zero, the difference between the groups is statistically significant at the .05 level.

The following example of comparing groups addresses the problem of determining whether the average mathematics scale score of group A is higher than that of group B. The sample estimates of the average scale scores and estimated standard errors are as follows:

Group	Average scale score	Standard error
A	218	0.9
B	216	1.1

The difference between the estimates of the average scale scores of groups A and B is 2 points (218 – 216). The standard error of this difference is

$$\sqrt{(0.9^2 + 1.1^2)} = 1.4$$

Thus, an approximately 95 percent confidence interval for

this difference is plus or minus 1.96 standard errors of the difference:

$$2 \pm 1.96 \times 1.4$$

$$2 \pm 2.7$$

$$(-0.7, 4.7)$$

The value zero is within the confidence interval; therefore, there is insufficient evidence to conclude that group A’s performance is statistically different from group B.

The procedure above is appropriate to use when it is reasonable to assume that the groups being compared have been independently sampled for the assessment.

Such an assumption is clearly warranted when comparing results for one state with another. This is the approach used for NAEP reports when comparisons involving independent groups are made. The assumption of independence is violated to some degree when comparing group results for the nation or a particular state (e.g., comparing national 2007 results for Black and White students), since these samples of students have been drawn from the same schools.

When the groups being compared do not share students (as is the case, for example, of comparing Black and White students), the impact of this violation of the independence assumption on the outcome of the statistical tests is assumed to be small, and NAEP, by convention, has, for computational convenience, routinely applied the procedures described above to those cases as well.

When making comparisons of results for groups that share a considerable proportion of students in common, it is not appropriate to ignore such dependencies. In such cases, NAEP has used procedures appropriate to comparing dependent groups. When the dependence in group results is due to the overlap in samples (e.g., when a subgroup is being compared to a total group), a simple modification of the usual standard error of the difference formula can be used. The formula for such cases is

$$SE^2_{Total-Subgroup} = \sqrt{(SE^2_{Total} + SE^2_{Subgroup} - 2pSE^2_{Subgroup})}$$

where p is the proportion of the total group contained in the subgroup. This formula was used for this report when a state was compared to the aggregate for the nation.

Conducting multiple tests

The procedures used to determine whether group differences in the long-term trend and main NAEP samples represent actual differences among the groups in the population and the certainty ascribed to intervals (e.g., a 95 percent confidence interval) are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. However, there are times when many different groups are being compared (i.e., multiple sets of confidence intervals are being analyzed).

For multiple comparisons, statistical theory indicates that the certainty associated with the entire set of comparisons is less than that attributable to each individual comparison from the set. To hold the significance level for the set of comparisons at a particular level (e.g., .05), the standard methods must be adjusted by multiple comparison procedures.³ The procedure used by NAEP is the Benjamini-Hochberg False Discovery Rate (FDR) procedure.⁴

Unlike other multiple comparison procedures that control the family-wise error rate (i.e., the probability of making even one false rejection in the set of comparisons), the FDR procedure controls the expected proportion of falsely rejected hypotheses. Furthermore, the FDR procedure used in NAEP is considered appropriately less conservative than family-wise procedures for large families of compari-

³ Miller, R.G. (1981). *Simultaneous Statistical Inference* (2nd ed.). New York: Springer-Verlang.

⁴ Benjamini, Y., and Hochberg, Y. (1995), op. cit.

sons.⁵ Therefore, the FDR procedure is more suitable for multiple comparisons in NAEP than are other procedures.

Statistical comparisons of NAEP scores from different assessment years are made using a multiple comparison procedure. However, in figures 9, 11, 21, and 23, comparisons of the size of the Black-White achievement gap for each state to the national gap are made using pairwise comparisons, where each state is compared to the nation one at a time. For this reason, the results shown in these four figures may not correspond to results obtained from the NAEP Online Data Tool, which currently does not permit pairwise comparisons for this type of gap analysis.

Cautions in interpretation

It is possible to examine NAEP performance results for groups of students defined by various background factors measured by NAEP, such as race. However, a relationship that exists between achievement and another variable does not reveal its underlying cause, which may be influenced by a number of other variables. Similarly, the assessments do not reflect the influence of unmeasured variables. The results are most useful when they are considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

⁵ Williams, V.S.L., Jones, L.V., and Tukey, J.W. (1994, December) *Controlling Error in Multiple Comparisons with Special Attention to the National Assessment of Educational Progress*. Research Triangle Park, NC: National Institute of Statistical Sciences.

Appendix B: Supplemental Tables

Table B-1. Administration of NAEP national and state mathematics assessments, by grade: Various years, 1990–2007

Grade	1990	1992	1994	1996	1998	2000	2002	2003	2005	2007
	National State	National State	National State	National State	National State	National State	National State	National State	National State	National State
4th grade	✓	✓		✓		✓		✓	✓	✓
8th grade	✓	✓		✓		✓		✓	✓	✓

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, Various years, 1990–2007 Mathematics Assessments.

Table B-2. Average national mathematics scale scores for all public school students at grades 4 and 8, by gender and eligibility for the National School Lunch Program: Various years, 1990–2007

	1990 ⁿ	1992 ⁿ	1996	2000	2003	2005	2007
All students							
Grade 4	212*	219*	222*	224*	234*	237*	239
Grade 8	262*	267*	269*	272*	276*	278*	280
Student Gender							
Grade 4							
Male	212*	220*	222*	225*	235*	238*	240
Female	211*	218*	222*	223*	233*	236*	238
Grade 8							
Male	262*	266*	270*	273*	277*	278*	281
Female	261*	267*	268*	271*	275*	277*	279
Student Eligibility for National School Lunch Program							
Grade 4							
Not eligible	—	—	‡	‡	244*	248*	249
Reduced-price lunch	—	—	‡	‡	230*	234*	236
Free lunch	—	—	‡	‡	220*	224*	226
Grade 8							
Not eligible	—	—	‡	‡	287*	288*	291
Reduced-price lunch	—	—	‡	‡	269*	270*	274
Free lunch	—	—	‡	‡	256*	260*	263

ⁿ Accommodations were not permitted for this assessment.

— Not available. Data were not collected prior to 1996.

‡ Reporting standards not met. Lunch eligibility data are not being reported in 1996 and 2000 because of the high percentage of students for whom information was not available.

* Significantly different (p<.05) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), Various years, 1990–2007 Mathematics Assessments.

Table B-3. Administration of NAEP national and state reading assessments, by grade: Various years, 1992–2007

Grade	1992		1994		1996		1998		2000		2002		2003		2005		2007	
	National	State	National	State	National	State	National	State	National	State	National	State	National	State	National	State	National	State
4th grade	✓	✓	✓	✓			✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
8th grade	✓		✓				✓	✓			✓	✓	✓	✓	✓	✓	✓	✓

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, Various years, 1992–2007 Reading Assessments.

Table B-4. Average national reading scale scores for all public school students at grades 4 and 8, by gender and eligibility for the National School Lunch Program: Various years, 1992–2007

	1992 ^a	1994 ^a	1998	2000	2002	2003	2005	2007
All students								
Grade 4	215*	212*	213*	211*	217*	216*	217*	220
Grade 8	258*	257*	261	—	263*	261	260*	261
Student Gender								
Grade 4								
Male	211*	207*	210*	206*	214*	213*	214*	216
Female	219*	218*	215*	217*	220*	220*	220*	223
Grade 8								
Male	251*	250*	253*	—	258*	256	255*	256
Female	264	265	268*	—	267*	267	266	266
Student Eligibility for National School Lunch Program								
Grade 4								
Not eligible	—	—	‡	‡	‡	229*	230*	232
Reduced-price lunch	—	—	‡	‡	‡	211*	212*	215
Free lunch	—	—	‡	‡	‡	199*	201*	203
Grade 8								
Not eligible	—	—	‡	—	‡	271	270*	271
Reduced-price lunch	—	—	‡	—	‡	256	254	255
Free lunch	—	—	‡	—	‡	243*	245	246

^a Accommodations were not permitted for this assessment.

— Not available. Data were not collected prior to 1996 or at grade 8 in 2000.

‡ Reporting standards not met. Lunch eligibility data are not being reported in 1998, 2000 and 2002 because of the high percentage of students for whom information was not available.

* Significantly different (p<.05) from 2007.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), Various years, 1992–2007 Reading Assessments.

Acknowledgments

This report was prepared under a project of the NAEP Education Statistics Services Institute (NAEP-ESSI) of the American Institutes for Research (AIR) in support of the National Center for Education Statistics (NCES), part of the Institute of Education Sciences of the U.S. Department of Education.

The authors are grateful to everyone who contributed to the design, development, review, and production of this report.

