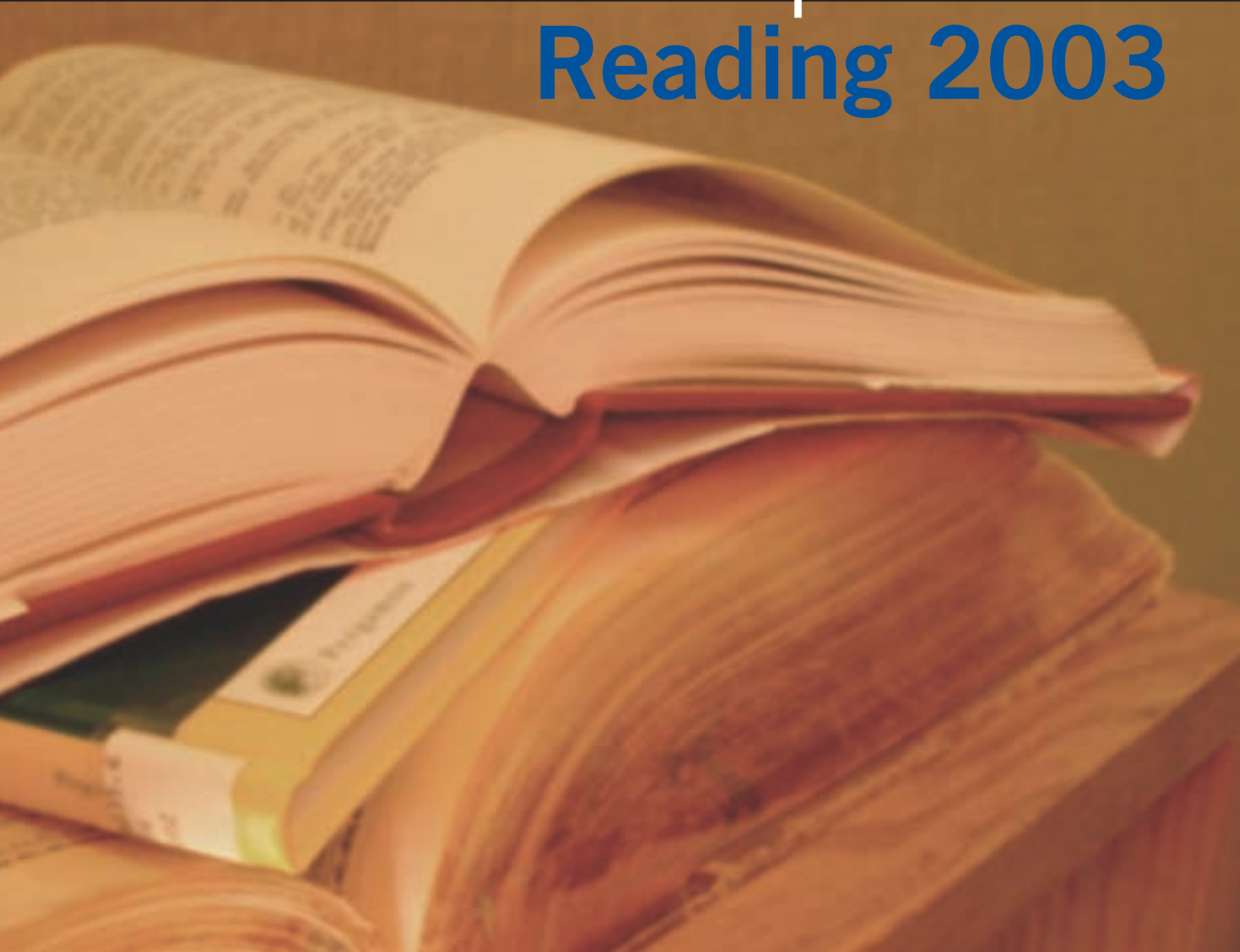




**U.S. Department of Education**  
Institute of Education Sciences  
NCES 2005-453

# The Nation's Report Card<sup>TM</sup> Reading 2003



The National Assessment of Educational Progress

## What is The Nation's Report Card™?

THE NATION'S REPORT CARD™, the National Assessment of Educational Progress (NAEP), is a nationally representative and continuing assessment of what America's students know and can do in various subject areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, history, geography, and other fields. By making objective information on student performance available to policymakers at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. Only information related to academic achievement is collected under this program. NAEP guarantees the privacy of individual students and their families.

NAEP is a congressionally mandated project of the National Center for Education Statistics within the Institute of Education Sciences of the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations.

In 1988, Congress established the National Assessment Governing Board (NAGB) to oversee and set policy for NAEP. The Board is responsible for: selecting the subject areas to be assessed; setting appropriate student achievement levels; developing assessment objectives and test specifications; developing a process for the review of the assessment; designing the assessment methodology; developing guidelines for reporting and disseminating NAEP results; developing standards and procedures for interstate, regional, and national comparisons; determining the appropriateness of all assessment items and ensuring the assessment items are free from bias and are secular, neutral, and non-ideological; taking actions to improve the form, content, use, and reporting of results of the National Assessment; and planning and executing the initial public release of National Assessment of Educational Progress reports.

## The National Assessment Governing Board

### **Darvin M. Winick, Chair**

President  
Winick & Associates  
Dickinson, Texas

### **Sheila M. Ford, Vice Chair**

Principal  
Horace Mann Elementary  
School  
Washington, D.C.

### **Francie Alexander**

Chief Academic Officer,  
Scholastic, Inc.  
Senior Vice President,  
Scholastic Education  
New York, New York

### **David J. Alukonis**

Chairman  
Hudson School Board  
Hudson, New Hampshire

### **Amanda P. Avallone**

Assistant Principal &  
Eighth-Grade Teacher  
Summit Middle School  
Boulder, Colorado

### **Honorable Jeb Bush**

Governor of Florida  
Tallahassee, Florida

### **Barbara Byrd-Bennett**

Chief Executive Officer  
Cleveland Municipal  
School District  
Cleveland, Ohio

### **Carl A. Cohn**

Clinical Professor  
Rossier School of  
Education  
University of  
Southern California  
Los Angeles, California

### **Shirley V. Dickson**

Educational Consultant  
Laguna Niguel, California

### **John Q. Easton**

Executive Director  
Consortium on Chicago  
School Reform  
Chicago, Illinois

### **Honorable Dwight Evans**

Member  
Pennsylvania House of  
Representatives  
Philadelphia, Pennsylvania

### **David W. Gordon**

Sacramento County  
Superintendent of Schools  
Sacramento County Office  
of Education  
Sacramento, California

### **Henry L. Johnson**

Superintendent of  
Education  
Mississippi Department  
of Education  
Jackson, Mississippi

### **Kathi M. King**

Twelfth-Grade Teacher  
Messalonskee High  
School  
Oakland, Maine

### **Honorable Keith King**

Member  
Colorado House of  
Representatives  
Colorado Springs, Colorado

### **Kim Kozbial-Hess**

Fourth-Grade Teacher  
Fall-Meyer Elementary  
School  
Toledo, Ohio

### **Andrew C. Porter**

Professor  
Leadership Policy  
and Organizations  
Vanderbilt University  
Nashville, Tennessee

### **Luis A. Ramos**

Community Relations  
Manager  
PPL Susquehanna  
Berwick, Pennsylvania

### **Mark D. Reckase**

Professor  
Measurement and  
Quantitative Methods  
Michigan State University  
East Lansing, Michigan

### **John H. Stevens**

Executive Director  
Texas Business and  
Education Coalition  
Austin, Texas

### **Mary Frances Taymans, SND**

Executive Director  
National Catholic  
Educational Association  
Washington, D.C.

### **Oscar A. Troncoso**

Principal  
Socorro High School  
Socorro Independent  
School District  
El Paso, Texas

### **Honorable Thomas J. Vilsack**

Governor of Iowa  
Des Moines, Iowa

### **Michael E. Ward**

Former State Superintendent  
of Public Instruction  
North Carolina Public  
Schools  
Jackson, Mississippi

### **Eileen L. Weiser**

Member, State Board  
of Education  
Michigan Department  
of Education  
Lansing, Michigan

### **Grover J. Whitehurst (Ex officio)**

Director  
Institute of Education  
Sciences  
U.S. Department of  
Education  
Washington, D.C.

---

### **Charles E. Smith**

Executive Director  
NAGB  
Washington, D.C.



U.S. Department of Education  
Institute of Education Sciences  
NCES 2005-453

# The Nation's Report Card™

# Reading

# 2003

Patricia L. Donahue  
Mary C. Daane  
Ying Jin

*In collaboration with*  
Hui Deng  
Kelvin Gregory  
Steven Isham  
Andreas Oranje  
Tatyana Petrovicheva  
Fred Schaeffer  
Jinming Zhang  
**Educational Testing Service**

Arnold A. Goldstein  
*Project Officer*  
**National Center for  
Education Statistics**

**U.S. Department of Education**

Margaret Spellings  
*Secretary*

**Institute of Education Sciences**

Grover J. Whitehurst  
*Director*

**National Center for Education Statistics**

Grover J. Whitehurst  
*Acting Commissioner*

---

**July 2005**

**SUGGESTED CITATION**

Donahue, P.L., Daane, M.C., and Jin, Y. (2005). *The Nation's Report Card: Reading 2003* (NCES 2005-453). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

**FOR MORE INFORMATION**

Content contact:  
Arnold Goldstein  
202-502-7344  
Arnold.Goldstein@ed.gov

To obtain single copies of this report, or for ordering information on other U.S. Department of Education products, call toll free 1-877-4ED-PUBS (877-433-7827), or write:

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

TTY/TDD 1-877-576-7734  
FAX 301-470-1244

Online ordering via the Internet: <http://www.edpubs.org>  
Copies also are available in alternate formats upon request.  
This report also is available on the World Wide Web: <http://nces.ed.gov/pubsearch>

The work upon which this publication is based was performed for the National Center for Education Statistics by Educational Testing Service, Pearson Educational Measurement, and Westat.

# T able of Contents

<b>Executive Summary</b> .....	<b>.xv</b>
--------------------------------	------------

## **Chapter 1**

<b>Introduction</b> .....	<b>1</b>
Overview of the 2003 National Assessment of Educational Progress in Reading .....	1
Framework for the 1992, 1994, 1998, 2000, 2002, and 2003 NAEP Reading Assessments .....	3
The 2003 NAEP Reading Assessment Instrument .....	5
Description of School and Student Samples .....	7
Reporting the Assessment Results .....	7
The Setting of Achievement Levels .....	8
Reading Achievement-Level Descriptions for Each Grade .....	9
Trial Status of Achievement Levels .....	12
Interpreting NAEP Results .....	14
Overview of the Remaining Report .....	15

## **Chapter 2**

<b>Average Reading Scale Score and Achievement-Level Results for the Nation and States</b> .....	<b>17</b>
Overview .....	17
National Reading Scale Score Results .....	18
National Reading Scale Scores by Percentile .....	19
National Reading Achievement-Level Results .....	20
Reading Results by Region of the Country .....	23
Reading Results for States and Other Jurisdictions .....	25
Reading Scale Score Results by State/Jurisdiction .....	25
Cross-State/Jurisdiction Reading Scale Score Comparisons .....	29
Reading Achievement-Level Results by State/Jurisdiction .....	32
Cross-State/Jurisdiction Reading Achievement-Level Comparisons .....	38

<b>Chapter 3</b>	
<b>Subgroup Results for the Nation and States</b>	<b>41</b>
Performance of Selected Subgroups for the Nation	42
Gender	42
Race/Ethnicity	46
Student Eligibility for Free/Reduced-Price School Lunch	53
Parents' Highest Level of Education	58
Type of School	62
Type of Location	66
Performance of Selected Subgroups by State	69
Gender	69
Race/Ethnicity	75
Student Eligibility for Free/Reduced-Price School Lunch	89
<b>Chapter 4</b>	
<b>Average Reading Scale Scores and Achievement-Level Results for Districts Participating in the Trial Urban District Assessment</b>	<b>95</b>
Scale Score Results for Urban Districts	96
Scale Scores by Percentiles for Urban Districts	98
Achievement-Level Results for Urban Districts	99
Performance of Selected Subgroups for Urban Districts	101
Gender	101
Race/Ethnicity	105
Student Eligibility for Free/Reduced-Price School Lunch	112
Highest Level of Parents' Education	118
<b>Chapter 5</b>	
<b>Sample Assessment Questions and Student Responses</b>	<b>121</b>
Grade 4 Sample Assessment Questions and Results	122
Grade 8 Sample Assessment Questions and Results	127
Maps of Selected Item Descriptions on the NAEP Reading Scale—Grades 4 and 8	132
<b>Appendix A</b>	
<b>Overview of Procedures Used for the NAEP 2003 Reading Assessment</b>	<b>135</b>
<b>Appendix B</b>	
<b>Subgroup Percentage Appendix</b>	<b>201</b>
<b>Appendix C</b>	
<b>State and Urban District Subgroup Appendix</b>	<b>221</b>
<b>Appendix D</b>	
<b>State- and District-Level Contextual Variables</b>	<b>253</b>
<b>Appendix E</b>	
<b>Sample Text from the NAEP 2003 Reading Assessment</b>	<b>259</b>
<b>Acknowledgments</b>	<b>265</b>

## Chapter 1: Tables and Figures

Figure 1.1	
Descriptions of the three contexts for reading in the NAEP reading assessment	4
Figure 1.2	
Descriptions of the four aspects of reading in the NAEP reading assessment	5
Table 1.1	
Percentage weighting of the “context for reading” subscales on the NAEP composite reading scale, grades 4 and 8	8
Figure 1.3	
Policy definitions of the three NAEP achievement levels	9
Figure 1.4	
Descriptions of NAEP reading achievement levels, grade 4	10
Figure 1.5	
Descriptions of NAEP reading achievement levels, grade 8	11

## Chapter 2: Tables and Figures

Figure 2.1	
Average reading scale scores, grades 4 and 8: 1992–2003	19
Figure 2.2	
Reading scale score percentiles, grades 4 and 8: 1992–2003	20
Figure 2.3	
Percentages of students at or above <i>Basic</i> and <i>Proficient</i> in reading, grades 4 and 8: 1992–2003	21
Table 2.1	
Percentages of students, by reading achievement level, grades 4 and 8: 1992–2003	22
Figure 2.4	
Map of regions of the country according to U.S. Census	23
Table 2.2	
Average reading scale scores, by region of the country, grades 4 and 8: 2003	24
Table 2.3	
Percentages of students, by reading achievement level and region of the country, grades 4 and 8: 2003	25
Table 2.4	
Average reading scale scores, grade 4 public schools: By state, 1992–2003	26
Table 2.5	
Average reading scale scores, grade 8 public schools: By state, 1998–2003	27
Figure 2.5	
Comparison of state and national public school average reading scale scores, grade 4: 2003	28
Figure 2.6	
Comparison of state and national public school average reading scale scores, grade 8: 2003	29
Figure 2.7	
Cross-state comparison of average reading scale scores, grade 4 public schools: 2003	30
Figure 2.8	
Cross-state comparison of average reading scale scores, grade 8 public schools: 2003	31

Figure 2.9	
Percentage of students within each reading achievement level, grade 4 public schools:	
By state, 2003	33
Figure 2.10	
Percentage of students within each reading achievement level, grade 8 public schools:	
By state, 2003	34
Table 2.6	
Percentage of students at or above <i>Proficient</i> in reading, grade 4 public schools:	
By state, 1992–2003	36
Table 2.7	
Percentage of students at or above <i>Proficient</i> in reading, grade 8 public schools:	
By state, 1998–2003	37
Figure 2.11	
Cross-state comparison of percentage of students at or above <i>Proficient</i> in reading, grade 4 public schools: 2003	39
Figure 2.12	
Cross-state comparison of percentage of students at or above <i>Proficient</i> in reading, grade 8 public schools: 2003	40

**Chapter 3: Tables and Figures**

Figure 3.1	
Average reading scale scores, by gender, grades 4 and 8: 1992–2003	43
Figure 3.2	
Gaps in average reading scale scores, by gender, grades 4 and 8: 1992–2003	44
Table 3.1	
Percentages of students, by reading achievement level and gender, grades 4 and 8: 1992–2003	45
Figure 3.3	
Average reading scale scores, by race/ethnicity, grades 4 and 8: 1992–2003	48
Figure 3.4	
Gaps in average reading scale scores, by race/ethnicity, grades 4 and 8: 1992–2003	49
Table 3.2	
Percentages of students, by reading achievement level and race/ethnicity, grades 4 and 8: 1992–2003	51
Figure 3.5	
Average reading scale scores, by students' eligibility for free/reduced-price school lunch, grades 4 and 8: 1998–2003	54
Figure 3.6	
Gaps in average reading scale scores, by students' eligibility for free/reduced-price school lunch, grades 4 and 8: 1998–2003	55
Table 3.3	
Percentages of students, by reading achievement level and eligibility for free/reduced-price school lunch, grades 4 and 8: 1998–2003	56
Table 3.4	
Average reading scale scores, by student eligibility for free/reduced-price school lunch and race/ethnicity, grades 4 and 8: 2003	58
Figure 3.7	
Average reading scale scores, by student-reported parents' highest level of education, grade 8: 1992–2003	59



Table 3.5	
Percentages of students, by reading achievement level and student-reported parents' highest level of education, grade 8: 1992–2003	61
Figure 3.8	
Average reading scale scores, by type of school, grades 4 and 8: 1992–2003	63
Table 3.6	
Percentages of students, by reading achievement level and type of school, grades 4 and 8: 1992–2003	64
Table 3.7	
Average reading scale scores, by student-reported parents' highest level of education and type of school, grade 8: 2003	66
Figure 3.9	
Average reading scale scores, by type of location, grades 4 and 8: 2000–2003	67
Table 3.8	
Percentages of students, by reading achievement level and type of location, grades 4 and 8: 2000–2003	68
Table 3.9	
Average reading scale scores, by gender, grade 4 public schools: By state, 1992–2003	70
Table 3.10	
Average reading scale scores, by gender, grade 8 public schools: By state, 1998–2003	71
Table 3.11	
Percentage of students at or above <i>Proficient</i> in reading, by gender, grade 4 public schools: By state, 1992–2003	73
Table 3.12	
Percentage of students at or above <i>Proficient</i> in reading, by gender, grade 8 public schools: By state, 1998–2003	74
Table 3.13	
Average reading scale scores, by race/ethnicity, grade 4 public schools: By state, 1992–2003	76
Table 3.14	
Average reading scale scores, by race/ethnicity, grade 8 public schools: By state, 1998–2003	79
Table 3.15	
Percentage of students at or above <i>Proficient</i> in reading, by race/ethnicity, grade 4 public schools: By state, 1992–2003	83
Table 3.16	
Percentage of students at or above <i>Proficient</i> in reading, by race/ethnicity, grade 8 public schools: By state, 1998–2003	86
Table 3.17	
Average reading scale scores, by student eligibility for free/reduced-price school lunch, grade 4 public schools: By state, 1998–2003	90
Table 3.18	
Average reading scale scores, by student eligibility for free/reduced-price school lunch, grade 8 public schools: By state, 1998–2003	91
Table 3.19	
Percentage of students at or above <i>Proficient</i> in reading, by student eligibility for free/reduced-price school lunch, grade 4 public schools: By state, 1998–2003	93
Table 3.20	
Percentage of students at or above <i>Proficient</i> in reading, by student eligibility for free/reduced-price school lunch, grade 8 public schools: By state, 1998–2003	94

## Chapter 4: Tables and Figures

Figure 4.1	
Average reading scale scores, grade 4 public schools: By urban district, 2002 and 2003 . . . . .	97
Figure 4.2	
Average reading scale scores, grade 8 public schools: By urban district, 2002 and 2003 . . . . .	97
Table 4.1	
Reading scale score percentiles, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . .	99
Table 4.2	
Percentages of students, by reading achievement level, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	100
Table 4.3	
Average reading scale scores, by gender, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	101
Figure 4.3	
Gaps in average reading scores, by gender, grades 4 and 8 public schools: By urban district, 2003 . . . . .	102
Table 4.4	
Percentages of students, by reading achievement level and gender, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	104
Table 4.5	
Average reading scale scores, by race/ethnicity, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	106
Figure 4.4	
Gaps in average reading scores, by race/ethnicity, grades 4 and 8 public schools: By urban district, 2003 . . . . .	107
Table 4.6	
Percentages of students, by reading achievement level and race/ethnicity, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	108
Table 4.7	
Average reading scale scores, by eligibility for free/reduced-price school lunch, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	113
Figure 4.5	
Gaps in average reading scores, by eligibility for free/reduced-price school lunch, grades 4 and 8 public schools: By urban district, 2003 . . . . .	114
Table 4.8	
Percentages of students, by reading achievement level and eligibility for free/reduced-price school lunch, grades 4 and 8 public schools: By urban district, 2002 and 2003 . . . . .	115
Table 4.9	
Average reading scale scores, by student-reported parents' highest level of education, grade 8 public schools: By urban district, 2002 and 2003 . . . . .	118
Table 4.10	
Percentages of students, by reading achievement level and student-reported parents' highest level of education, grade 8 public schools: By urban district, 2002 and 2003 . . . . .	119

## Chapter 5: Tables and Figures

Table 5.1 Percentage scored correct for multiple-choice sample question 1, by achievement-level range, grade 4: 2003 .....	122
Table 5.2 Percentage scored correct for multiple-choice sample question 2, by achievement-level range, grade 4: 2003 .....	123
Table 5.3 Percentage scored as “Acceptable” for short constructed-response sample question 3, by achievement-level range, grade 4: 2003 .....	124
Table 5.4a Percentage scored as “Essential” or better for extended constructed-response sample question 4, by achievement-level range, grade 4: 2003 .....	125
Table 5.4b Percentage scored as “Extensive” for extended constructed-response sample question 4, by achievement-level range, grade 4: 2003 .....	126
Table 5.5 Percentage scored correct for multiple-choice sample question 5, by achievement-level range, grade 8: 2003 .....	127
Table 5.6 Percentage scored correct for multiple-choice sample question 6, by achievement-level range, grade 8: 2003 .....	128
Table 5.7 Percentage scored as “Full Comprehension” for short constructed-response sample question 7, by achievement-level range, grade 8: 2003 .....	129
Table 5.8a Percentage scored as “Essential” or better for extended constructed-response sample question 8, by achievement-level range, grade 8: 2003 .....	130
Table 5.8b Percentage scored as “Extensive” for extended constructed-response sample question 8, by achievement-level range, grade 8: 2003 .....	131
Figure 5.1 Map of selected item descriptions on the NAEP reading scale, grade 4: 2003 .....	133
Figure 5.2 Map of selected item descriptions on the NAEP reading scale, grade 8: 2003 .....	134

## Appendix A: Tables and Figures

Figure A.1 Sample NAEP questions, by aspects of reading and contexts for reading specified in the reading framework .....	137
Table A.1 Target and actual percentage distribution of questions, by context for reading, grades 4 and 8: 2003 .....	138
Table A.2 Target and actual percentage distribution of student time, by aspect of reading, grades 4 and 8: 2003 .....	138
Table A.3 Number of students assessed, by sample type, special needs status, and accommodation option, grades 4 and 8 public and nonpublic schools: 1992–2003 .....	142
Table A.4 National and state sample sizes and target populations, grades 4 and 8: 2003 .....	143
Table A.5 District sample sizes and target populations, grades 4 and 8: 2003 .....	144
Table A.6 National school and student participation rates, by type of school, grades 4 and 8: 2003 .....	146
Table A.7 School and student participation rates, grade 4 public schools: By state, 2003 .....	147
Table A.8 School and student participation rates, grade 8 public schools: By state, 2003 .....	148
Table A.9 Weighted school and student participation rates, grades 4 and 8 public schools: By urban district, 2003 .....	149
Table A.10 Students with disabilities and/or limited-English-proficient students identified, excluded, and assessed, when accommodations were not permitted, grades 4 and 8 public and nonpublic schools: 1992–2000 .....	153
Table A.11 Percentage of students with disabilities and/or limited-English-proficient students identified, excluded, and assessed, when accommodations were not permitted, grade 4 public schools: By state, 1992–1998 .....	154
Table A.12 Percentage of students with disabilities and/or limited-English-proficient students identified, excluded, and assessed, when accommodations were not permitted, grade 8 public schools: By state, 1998 .....	155
Table A.13 Students with disabilities and/or limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grades 4 and 8 public and nonpublic schools: 1998–2003 .....	156
Table A.14 Percentage of students with disabilities and/or limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grade 4 public schools: By state, 1998–2003 .....	157

Table A.15	
Percentage of students with disabilities identified, excluded, and assessed, when accommodations were permitted, grade 4 public schools: By state, 1998–2003	160
Table A.16	
Percentage of limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grade 4 public schools: By state, 1998–2003	163
Table A.17	
Percentage of students with disabilities and/or limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grade 8 public schools: By state, 1998–2003	166
Table A.18	
Percentage of students with disabilities identified, excluded, and assessed, when accommodations were permitted, grade 8 public schools: By state, 1998–2003	169
Table A.19	
Percentage of limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grade 8 public schools: By state, 1998–2003	172
Table A.20	
Percentage of students with disabilities and limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grade 4 public schools: By urban district, 2002 and 2003	175
Table A.21	
Percentage of students with disabilities and limited-English-proficient students identified, excluded, and assessed, when accommodations were permitted, grade 8 public schools: By urban district, 2002 and 2003	176
Table A.22	
Students with disabilities and/or limited-English-proficient students assessed with accommodations, by type of primary accommodation, grades 4 and 8 public and nonpublic schools: 1998–2003	179
Table A.23	
Average reading scale scores and standard errors, grades 4 and 8: 1992–2003	187
Table A.24	
Percentage of students and standard errors, by reading achievement level, grades 4 and 8: 1992–2003	188
Table A.25	
Average reading scale scores and standard errors, by race/ethnicity and eligibility for free/reduced-price school lunch, grades 4 and 8: 2003	188
Table A.26	
Average reading scale scores and standard errors, grade 8 public schools: By state, 1998–2003	189
Table A.27	
Percentage of students at or above <i>Proficient</i> and standard errors, by race/ethnicity, grade 8 public schools: By state, 1998–2003	190
Table A.28	
Example of False Discovery Rate comparisons of average scale scores for different groups of students	195
Figure A.2	
States within regions of the country defined by the U.S. Census Bureau	198

## Appendix B: Tables

Table B.1	
Weighted percentage of students, by region of the country, grades 4 and 8: 2003	202
Table B.2	
Weighted percentage of students, by gender, grades 4 and 8: 1992-2003	202
Table B.3	
Weighted percentage of students, by race/ethnicity, grades 4 and 8: 1992-2003	203
Table B.4	
Weighted percentage of students, by eligibility for free/reduced-price school lunch, grades 4 and 8: 1998-2003	203
Table B.5	
Weighted percentages of students, by eligibility for free/reduced-price school lunch and race/ethnicity, grades 4 and 8: 2003	204
Table B.6	
Weighted percentage of students, by student-reported parents' highest level of education, grade 8: 1992-2003	204
Table B.7	
Weighted percentage of students, by type of school, grades 4 and 8: 1992-2003	205
Table B.8	
Weighted percentages of students, by parents' highest level of education and type of school, grade 8: 2003	205
Table B.9	
Weighted percentage of students, by type of location, grades 4 and 8: 2000-2003	205
Table B.10	
Weighted percentage of students, by gender, grade 4: By state, 1992-2003	206
Table B.11	
Weighted percentage of students, by gender, grade 8: By state, 1998-2003	207
Table B.12	
Weighted percentage of students, by race/ethnicity, grade 4: By state, 1992-2003	208
Table B.13	
Weighted percentage of students, by race/ethnicity, grade 8: By state, 1998-2003	211
Table B.14	
Weighted percentage of students, by eligibility for free/reduced-price school lunch, grade 4: By state, 1998-2003	214
Table B.15	
Weighted percentage of students, by eligibility for free/reduced-price school lunch, grade 8: By state, 1998-2003	215
Table B.16	
Weighted percentage of students, by gender, grades 4 and 8 public schools: By urban district, 2002 and 2003	216
Table B.17	
Weighted percentage of students, by race/ethnicity, grades 4 and 8 public schools: By urban district, 2002 and 2003	217
Table B.18	
Weighted percentage of students, by eligibility for free/reduced-price school lunch, grades 4 and 8 public schools: By urban district, 2002 and 2003	218
Table B.19	
Weighted percentage of students, by student-reported parents' highest level of education, grade 8 public schools: By urban district, 2002 and 2003	219

## Appendix C: Tables

Table C.1	
Gaps in average reading scale scores, by gender, grade 4 public schools: By state, 1992-2003	222
Table C.2	
Gaps in average reading scale scores, by gender, grade 8 public schools: By state, 1998-2003	223
Table C.3	
Percentages of students, by gender and reading achievement level, grade 4 public schools: By state, 2003	224
Table C.4	
Percentages of students, by gender and reading achievement level, grade 8 public schools: By state, 2003	225
Table C.5	
Percentage of students at or above <i>Basic</i> in reading, by gender, grade 4 public schools: By state, 1992-2003	226
Table C.6	
Percentage of students at or above <i>Basic</i> in reading, by gender, grade 8 public schools: By state, 1998-2003	227
Table C.7	
Gaps in average reading scale scores, by race/ethnicity, grade 4 public schools: By state, 1992-2003	228
Table C.8	
Gaps in average reading scale scores, by race/ethnicity, grade 8 public schools: By state, 1998-2003	229
Table C.9	
Percentages of students, by race/ethnicity and reading achievement level, grade 4 public schools: By state, 2003	230
Table C.10	
Percentages of students, by race/ethnicity and reading achievement level, grade 8 public schools: By state, 2003	233
Table C.11	
Percentage of students at or above <i>Basic</i> in reading, by race/ethnicity, grade 4 public schools: By state, 1992-2003	236
Table C.12	
Percentage of students at or above <i>Basic</i> in reading, by race/ethnicity, grade 8 public schools: By state, 1998-2003	239
Table C.13	
Percentages of students, by eligibility for free/reduced-price school lunch and reading achievement level, grade 4 public schools: By state, 2003	242
Table C.14	
Percentages of students, by eligibility for free/reduced-price school lunch and reading achievement level, grade 8 public schools: By state, 2003	243

Table C.15	Percentage of students at or above <i>Basic</i> in reading, by student eligibility for free/reduced-price school lunch, grade 4 public schools: By state, 1998–2003	244
Table C.16	Percentage of students at or above <i>Basic</i> in reading, by student eligibility for free/reduced-price school lunch, grade 8 public schools: By state, 1998–2003	245
Table C.17	Average reading scale scores and achievement-level results, by students with and without disabilities and limited English proficiency, grade 4 public schools: By state, 2003	246
Table C.18	Average reading scale scores and achievement-level results, by students with and without disabilities and limited English proficiency, grade 8 public schools: By state, 2003	248
Table C.19	Average reading scale score and achievement-level results, by students with disabilities or limited-English-proficient students, grade 4 public schools: By urban district, 2003	250
Table C.20	Average reading scale score and achievement-level results, by students with disabilities or limited-English-proficient students, grade 8 public schools: By urban district, 2003	251

#### Appendix D: Tables

Table D.1	Population and public school enrollment, from non-NAEP sources: By state, April 2000 and fall 2000	254
Table D.2	Poverty status of school-age children and children served under Individuals with Disabilities Education Act and Chapter 1, from non-NAEP sources: By state, 2001 and school years 1990–1991 through 2000–2001	255
Table D.3	Expenditure per pupil, average teacher salary, and pupil/teacher ratio in public schools, from non-NAEP sources: By state, school years 1999–2000, 2001–2002, and fall 2000	256
Table D.4	Enrollment, expenditure per pupil, and pupil/teacher ratio in public schools, from non-NAEP sources: By urban district, fall 2000 and school year 1999–2000	257



# E

## xecutive Summary

The National Assessment of Educational Progress (NAEP) is an ongoing, nationally representative sample survey of student achievement in core subject areas. Authorized by Congress and administered by the National Center for Education Statistics (NCES) within the Institute of Education Sciences of the U.S. Department of Education, NAEP regularly reports to the public on the educational progress of fourth-, eighth-, and twelfth-grade students.

This report presents results of the NAEP 2003 fourth- and eighth-grade reading assessments for the nation, for regions of the country, for participating states and other jurisdictions, and for participating urban districts. Assessment results are described in terms of students' average reading score on a 0–500 scale and in terms of the percentage of students attaining each of three achievement levels: *Basic*, *Proficient*, and *Advanced*. National and district-level scores at different percentiles on the scale (indicating the percentage of students whose scores fell below a particular point) are also discussed.

The achievement levels are performance standards adopted by the National Assessment Governing Board (NAGB) as part of its statutory responsibilities. The achievement levels are a collective judgment of what students should know and be able to do for each grade tested. As provided by law, NCES, upon review of a congressionally mandated evaluation of NAEP,

determined that the achievement levels are to be used on a trial basis and should be interpreted with caution. However, both NCES and NAGB believe these performance standards are useful for understanding trends in student achievement. They have been widely used by national and state officials and others as a common yardstick of academic performance.

Approximately 188,000 fourth-graders from 7,500 schools and 155,000 eighth-graders from 6,100 schools were assessed. The national results reflect the performance of students attending both public and nonpublic schools, while the results for participating states and other jurisdictions, and for urban districts, reflect the performance of students attending public schools. In addition to providing average scores and achievement-level percentages in reading for the nation, states and other jurisdictions, and urban districts, this report provides results for subgroups of students defined by various background characteristics.

A summary of major findings from the NAEP 2003 Reading Assessment is presented on the following pages. Comparisons are made to results from previous years in which the assessment was administered. In addition to the 2003 results, national results are reported from the 1992, 1994, 1998, 2000 (fourth grade only), and 2002 assessments. Results for participating states and other jurisdictions are also reported from the 1992, 1994, 1998, and 2002 assessments at grade 4 and from the 1998 and 2002 assessments at grade 8. Results for participating urban districts are reported for 2002 and 2003. The more recent results (those from 1998 or later) are based on

administration procedures in which testing accommodations were permitted for students with disabilities and limited-English-proficient students. Accommodations were not permitted in earlier assessments. Comparisons between results from 2003 and those from assessment years in which both types of administration procedures were used (1998 at both grades and 2000 at grade 4 only) are discussed in this executive summary based on the results when accommodations were permitted. Changes in student performance across years or differences between groups of students in 2003 are discussed only if they have been determined to be statistically significant at the .05 level.

## **Overall Reading Results for the Nation, Regions of the Country, and States and Other Jurisdictions**

### **Reading Results for the Nation**

#### **At grade 4**

- No measurable difference was detected between the fourth-grade average score in 2003 and the score in 1992.
- The score at the 75th percentile for fourth-graders was higher in 2003 than in 1992, indicating improvement among higher-performing students.
- The percentage of fourth-graders performing at or above *Proficient* was higher in 2003 than in 1992.

#### **At grade 8**

- The average eighth-grade reading score decreased by one point between 2002 and 2003; however, the score in 2003 was higher than that in 1992.

- Scores decreased from 2002 to 2003 among lower-performing eighth-graders at the 10th and 25th percentiles; however, scores at the 10th, 25th, 50th, and 75th percentiles were higher in 2003 than in 1992.
- The percentage of eighth-graders at or above *Proficient* was higher in 2003 than in 1992. The percentage of students at or above *Basic* decreased by one point between 2002 and 2003, but was higher in 2003 than in 1992.

### Reading Results for Regions of the Country

Prior to 2003, NAEP results were reported for four NAEP-defined regions of the nation: Northeast, Southeast, Central, and West. As of 2003, to align NAEP with other federal data collections, NAEP analysis and reports have used the U.S. Census Bureau's definition of "region." The four regions defined by the U.S. Census Bureau are Northeast, South, Midwest, and West.

#### At grade 4

- The average fourth-grade reading score in 2003 was higher for students in the Northeast than in the Midwest, South, and West. In the Midwest, the average score was higher than in the South and West, and the average score was higher for students in the South than for students in the West.
- The percentages of fourth-graders performing at or above the *Basic* and *Proficient* levels in 2003 were higher in the Northeast than in the Midwest, South, and West. Higher percentages of students performed at or above the *Basic* and *Proficient* levels in the Midwest than in the South and the West, and higher percentages of students performed at or above the *Basic* and *Proficient* levels in the South than in the West.

#### At grade 8

- In 2003, the average eighth-grade reading scores were higher in the Northeast and Midwest than in the South and West, and the average score was higher in the South than in the West.
- Higher percentages of eighth-graders performed at or above the *Basic* and *Proficient* levels in 2003 in the Northeast and Midwest than in the South and West. In the South, a higher percentage of students performed at or above the *Basic* level than in the West.

### Reading Results for the States and Other Jurisdictions

Results from the 2003 assessment are reported for fourth- and eighth-grade students attending public schools in 50 states and 3 other jurisdictions that participated in the assessment. (Throughout this summary, the term "jurisdiction" is used to refer to the states, the District of Columbia, and the Department of Defense schools that participated in the NAEP reading assessments.)

#### At grade 4

- Of the 42 jurisdictions that participated in both the 1992 and 2003 fourth-grade assessments, 13 showed increases and 5 showed declines in average scores.
- The percentage of fourth-graders at or above *Proficient* increased in 17 of the 42 jurisdictions that participated in both the 1992 and 2003 assessments.
- Connecticut, Massachusetts, New Hampshire, New Jersey, and Vermont were among the jurisdictions with the highest average reading scores at grade 4.

### **At grade 8**

- Of the 39 jurisdictions that participated in the eighth-grade assessment in 1998 (when accommodations were permitted) and 2003, 8 showed increases and 7 showed declines in average scores.
- Between 1998 (when accommodations were permitted) and 2003, the percentage of eighth-graders performing at or above *Proficient* increased in 5 of the 39 jurisdictions that participated in both years, and declined in one.
- Department of Defense overseas schools, Massachusetts, New Hampshire, and Vermont were among the jurisdictions with the highest average reading scores at grade 8.

### **Reading Results for Student Subgroups in the Nation and in the States and Other Jurisdictions**

In addition to reporting overall results, NAEP reports on the performance of various subgroups of students. In interpreting these data, readers are reminded that the relationship between contextual variables and student performance is not necessarily causal. There are many other educational, cultural, and social factors that play a role in student achievement in a particular subject area.

### **National Results**

#### **Gender**

- At grade 4, there was no measurable difference detected in the average reading scores for male or female students from 1992 to 2003. At grade 8, the average score for male students in 2003 was higher than in 1992, and lower than in 2002.
- In 2003, female students outperformed male students by 7 points on average at grade 4 and by 11 points on average at grade 8. The fourth- and eighth-grade reading score gaps between male and female students showed no measurable change from 1992 to 2003.
- The percentages of male and female fourth-graders performing at or above *Proficient* showed no measurable change from 1992 to 2003. The percentage of male eighth-graders at or above *Proficient* was higher in 2003 than in 1992.

## Race/Ethnicity

- At grade 4, White students and Asian/Pacific Islander students scored higher on average in 2003 than Black, Hispanic, and American Indian/Alaska Native students. White students also scored higher on average than Asian/Pacific Islander students, and Hispanic students scored higher on average than Black students. At grade 8, White and Asian/Pacific Islander students had higher average scores in 2003 than Black, Hispanic, and American Indian/Alaska Native students.
- The average scores for White, Black, and Asian/Pacific Islander fourth-graders were higher in 2003 than in 1992. The average scores for White, Black, and Hispanic eighth-graders were higher in 2003 than in 1992.
- At both grades 4 and 8, the average score gap between White students and Black students and between White students and Hispanic students showed no measurable change from 1992 to 2003.
- The percentages of White, Black, and Asian/Pacific Islander fourth-graders performing at or above *Proficient* were higher in 2003 than in 1992. At grade 8, the percentages of White students and Black students performing at or above *Proficient* were higher in 2003 than in 1992.

## Eligibility for Free/Reduced-Price School Lunch

NAEP collects data on students' eligibility for free/reduced-price lunch as an indicator of family economic status. Eligibility for free/reduced-price lunch is determined by students' family income in relation to the federally established poverty level. The reading results are reported for students classified by their eligibility from 1998 on.

- In 2003, both fourth- and eighth-grade students who were eligible for free/reduced-price lunch scored lower on average than students who were not eligible.
- The average reading score for fourth-graders was higher in 2003 than in 1998 both for students who were eligible for free/reduced-price lunch and for those who were not eligible. The average score for eighth-graders who were eligible showed a decrease between 2002 and 2003 but showed no measurable difference between 1998 and 2003.
- For fourth-graders who were eligible for free/reduced-price lunch, the percentage at or above *Proficient* was higher in 2003 than in 1998.

### **Parents' Level of Education**

Eighth-grade students who participated in the NAEP reading assessment were asked to indicate the highest level of education completed by each parent. Results are reported based on the highest level of education for either parent. Information about parental education was not collected at grade 4.

- Overall, in 2003 there was a positive relationship between student-reported parental education and student achievement: the higher the parental education level, the higher the average reading score.
- The average score for eighth-grade students was lower in 2003 than in 2002 for students who reported that at least one parent had graduated from high school. The average score increased between 1992 and 2003 for students who reported that at least one parent had graduated from high school, and for students who reported that at least one parent had graduated from college.

### **Type of School**

The schools that participate in the NAEP assessment are classified as either public or nonpublic. A further distinction is then made between nonpublic schools that are Catholic schools and those that are some other type of nonpublic school.

- Performance results in 2003 show that, at both grades 4 and 8, students who attended nonpublic schools had a

higher average reading score than students who attended public schools.

- The average fourth-grade reading score for Catholic school students increased between 1992 and 2003. The average eighth-grade score was also higher in 2003 than in 1992 for Catholic school students. The average score for students in public schools declined between 2002 and 2003; however, the average public school score was higher in 2003 than that in 1992.
- The percentage of fourth-grade Catholic school students performing at or above *Proficient* was higher in 2003 than in 1992.

### **Type of Location**

The schools from which NAEP draws its samples of students are classified according to their type of location (central city, rural/small town, or urban fringe/large town). The methods used to identify the type of school location in 2000 (at grade 4), 2002, and 2003 were different from those used for prior assessment years; therefore, only the data from the 2000, 2002, and 2003 assessments are reported.

- In 2003, fourth- and eighth-graders in urban fringe/large town and rural/small town locations had higher average scores than students in central city locations, and students in urban fringe/large town locations scored higher on average than those in rural/small town locations.

- The average reading scores for fourth-graders in central city and urban fringe/large town locations were higher in 2003 than in 2000. The average score for eighth-graders in rural/small town locations declined between 2002 and 2003.
- In 2003, higher percentages of fourth- and eighth-graders performed at or above *Proficient* in urban fringe/large town and rural/small town locations than in central city locations.

### State and Other Jurisdiction Results

#### Gender

- In 2003, female students scored higher on average than male students in all 53 of the jurisdictions that participated at grades 4 and 8.
- Among the 42 jurisdictions that participated in both the 1992 and 2003 fourth-grade reading assessments, 10 showed increases in the average score for both male and female students. New Mexico and Oklahoma showed decreases for both male and female students.
- Among the 39 jurisdictions that participated in both the 1998 and 2003 eighth-grade reading assessments, Delaware and Missouri showed average score increases for both male and female students, and Arizona, Nevada, and New Mexico showed decreases for both male and female students.

#### Race/Ethnicity

- The average fourth-grade reading score was higher in 2003 than in 1992 for White students in 19 jurisdictions, for Black students in 8 jurisdictions, for Hispanic students in 5 jurisdictions, and for Asian/Pacific Islander students in 4 jurisdictions. The average score declined between 1992 and 2003 for Black students in Iowa and for American Indian/Alaska Native students in New Mexico. Average score increases were observed between 1992 and 2003 for three or more racial/ethnic subgroups in California, Florida, Maryland, and New York.
- The average eighth-grade reading score was higher in 2003 than in 1998 for White students in six jurisdictions, Black students in Delaware, and Asian/Pacific Islander students in Hawaii and Minnesota. A decrease in the average score was detected between 1998 and 2003 for White students in Maine, Black students in Oklahoma, and Hispanic students in New Mexico.

#### Eligibility for Free/Reduced-Price School Lunch

- The average fourth-grade reading score was higher in 2003 than in 1998 both for students who were eligible and students who were not eligible for free/reduced-price school lunch in 11 jurisdictions, for eligible students in 5 jurisdictions, and for students who were not eligible in 5 jurisdictions. In the District of Columbia, the average score increased for eligible students and decreased for students who were not eligible.

- The average eighth-grade reading score was higher in 2003 than in 1998 both for students who were eligible and students who were not eligible in Delaware and Missouri. Average scores were lower in 2003 than in 1998 for eligible students in New Mexico and Oklahoma, and for students who were not eligible in Nevada.

### Urban District Results

The 2002 Trial Urban District Assessment (TUDA) included five urban public school districts (Atlanta City School District, City of Chicago School District 299, Houston Independent School District, Los Angeles Unified School District, and New York City Public Schools) plus the District of Columbia. The same districts, plus four more (Boston Public School District, Charlotte-Mecklenburg Schools, Cleveland Municipal School District, and San Diego City Unified School District), participated in the 2003 TUDA.

### Overall Reading Results for the Urban Districts

#### At grade 4

- The average fourth-grade reading score in 9 of the 10 districts participating in 2003 was lower than the national public school score. Average fourth-grade reading scores in Atlanta, Chicago, Cleveland, the District of Columbia, and Los Angeles were lower than the average score for large central cities. Average scores in Charlotte and New York were higher than the large central city score.

- When compared to fourth-grade public school students in large central cities, scores at the 10th percentile were higher in Boston, Charlotte, Houston, and New York; scores at the 25th percentile were higher in Charlotte, Houston, and New York; scores at the 50th percentile were higher in Charlotte and New York; and scores at the 75th and 90th percentiles were higher in Charlotte.

- The percentage of fourth-graders at or above *Proficient* in 2003 was lower in 9 of the 10 districts when compared to the nation. In Charlotte, the percentage of students at or above *Proficient* was higher than the percentage for large central cities.

#### At grade 8

- Average eighth-grade reading scores in 9 of the 10 districts that participated in 2003 were lower than the national average score. Students in Atlanta, Cleveland, the District of Columbia, Houston, and Los Angeles scored lower on average than students in large central cities. Students in Boston and Charlotte had higher average scores than students in large central cities.
- In comparison to the scores for eighth-grade public school students in large central cities, scores at the 10th and 25th percentiles were higher in Charlotte, scores at the 50th percentile were higher in Charlotte and New York, and scores at the 75th and 90th percentiles were higher in Boston and Charlotte.



- In 2003, the percentage of eighth-graders at or above *Proficient* was lower in 9 of the 10 districts as compared to the nation. The percentages at or above *Proficient* were higher in Boston and Charlotte than in large central cities.

## Results for Student Subgroups in Urban Districts

### Gender

- At grade 4, the average score for female students in Charlotte was higher than that in the nation. Reading scores for male and female students in Charlotte were both higher on average than for male and female students in large central cities. Female students in New York had higher average scores than female students in large central cities.
- At grade 8, male and female students in all the districts that participated in 2003, except Charlotte, had lower average scores than their counterparts in the nation. Average scores for both male and female students in Charlotte were higher than for their counterparts in large central cities.

### Race/Ethnicity

- At grade 4, the average scores in 2003 for White students in Atlanta, Charlotte, the District of Columbia, and Houston; Black students in Charlotte and Houston; and Hispanic students in New York were higher than the corresponding scores in the nation and large central cities. The average scores for White students in Cleveland and Los Angeles; Black students in the District of Columbia; and Hispanic students in the District of Columbia and Los Angeles were lower than the corresponding scores in the nation and large central cities.
- In 2003 at grade 8, average reading scores for both White and Black students in Charlotte, and Hispanic students in Chicago were higher than comparable scores in the nation and large central cities. The average scores for White students in Cleveland; Black students in Atlanta, the District of Columbia, and Los Angeles; and Hispanic students in Los Angeles were lower than the scores in the nation and large central cities.

### ***Eligibility for Free/Reduced-Price Lunch***

- At grade 4, average scores in 2003 were higher for eligible students in New York and for students who were not eligible in Charlotte and New York compared to the corresponding scores in the nation and large central cities. Eligible students in Atlanta, the District of Columbia, and Los Angeles, and students who were not eligible in the District of Columbia, scored lower on average than comparable groups of students in the nation and large central cities.
- At grade 8, eligible students in Boston, Chicago, and New York, and students who were not eligible in Charlotte and New York scored higher on average than their counterparts in large central cities. Eligible students in Atlanta, the District of Columbia, and Los Angeles, and students who were not eligible in Atlanta, the District of Columbia, Houston, and Los Angeles, scored lower on average than their counterparts in the nation and large central cities.

### ***Parents' Level of Education***

- In 2003, the average score for eighth-grade students who indicated that a parent had graduated from college was lower in Atlanta, Chicago, Cleveland, the District of Columbia, and Los Angeles than the average score for students in the same parental education category in public schools in the nation and large central cities. The average score for students who reported that a parent graduated from college was higher in Charlotte than for comparable students in large central cities.

# 1

## Introduction

The importance of being able to read has long been acknowledged as the foundation for learning and as essential for participation in society. This report presents major results from the National Assessment of Educational Progress (NAEP) 2003 reading assessment of the nation's fourth- and eighth-grade students. Results are presented for the nation overall, for the 53 states and other jurisdictions that participated in the 2003 assessment, and for the 9 districts that participated in the Trial Urban District Assessment (TUDA). The results reported here are intended to inform educators, policymakers, parents, and the general public about students' progress in reading.

### **Overview of the 2003 National Assessment of Educational Progress in Reading**

For more than 30 years, NAEP has regularly collected, analyzed, and reported valid and reliable information about what students know and can do in a variety of subject areas. As authorized by the U.S. Congress, NAEP assesses representative national samples of fourth-, eighth-, and twelfth-grade students. Since 1992, NAEP has also assessed representative samples of fourth- and eighth-grade students in states and other jurisdictions that participate in the NAEP state-by-state assessments.

NAEP is administered and overseen by the National Center for Education Statistics (NCES), within the U.S. Department of Education's Institute of Education Sciences.

The content of all NAEP assessments is determined by subject-area frameworks that are developed by the National Assessment Governing Board (NAGB) in a comprehensive process involving a broad spectrum of interested parties, including teachers, curriculum specialists, subject-matter specialists, school administrators, parents, and members of the general public. The framework for the NAEP 2003 reading assessment, while updated and expanded, is in essence the same framework that has guided development of the NAEP reading assessments since 1992.

This report describes the results of the NAEP 2003 reading assessment at grades 4 and 8. National results for 2003 are compared to those from 1992, 1994, 1998, 2000, and 2002 at grade 4, and 1992, 1994, 1998, and 2002 at grade 8. Comparisons across assessment years are possible because the assessments were developed under the same basic framework and shared a common set of reading questions.

Using the same test as that used nationally, state-level assessments were conducted at grade 4 in 1992, 1994, 1998, 2002, and 2003. At grade 8, state-level assessments were conducted in 1998, 2002, and 2003. District-level results are presented for 9 districts in 2003 and for 5 districts in 2002.

Prior to 1998, administration procedures for NAEP reading assessments did not permit the use of accommodations for special needs students who could not

participate without them (e.g., extra time; individual rather than group administration). For the 1998 assessment, however, administration procedures were introduced that allowed the use of accommodations by students with disabilities (SD) and limited-English-proficient (LEP) students (see appendix A). A split-sample design was used in 1998 at all three grades (and again in 2000 at grade 4) so that both administration procedures could be used during the same assessment, but with different samples of students. This made it possible to report trends in students' reading achievement across all the assessment years and, at the same time, examine the effects of including students assessed with accommodations in overall assessment results. Based on an examination of how permitting accommodations affected overall population results, it was decided that, beginning with the 2002 assessment, NAEP would use only one set of procedures—permitting the use of accommodations.

During the period in which accommodations were not permitted, special-needs students could only be included in the assessment if it was determined by school staff that they could be assessed meaningfully without accommodations. The change in administration procedures makes it possible for more students to be included in the assessments; however, it also represents an important altering of procedures from previous assessments. (See the section on Students with Disabilities and/or Limited-English-Proficient students in appendix A for a more detailed discussion.) The reader is encouraged to consider the difference in accommodation procedures when interpreting comparisons between the two sets of results.

The charts and tables throughout this report distinguish between results from assessment years in which accommodations were not permitted and results from assessment years in which accommodations were permitted. In the tables and charts that display results across assessment years, all previous assessment results that were found to be significantly different (at the .05 level) from the 2003 results are marked with an asterisk (\*). Two sets of results are presented for assessment years in which both administration procedures were used (accommodations not permitted and accommodations permitted). Both sets of results may be notated, if found to be significantly different from 2003. The text that accompanies these tables and charts indicates which previous assessment results were significantly different from 2003. Comparisons between the 2003 results, when accommodations were permitted, and the 1992 and 1994 results, when they were not permitted, are discussed in the text. However, for assessment years with both accommodations-not-permitted results and accommodations-permitted results, the text describes comparisons only between the accommodations-permitted results and 2003.

### **Framework for the 1992, 1994, 1998, 2000, 2002, and 2003 NAEP Reading Assessments**

The reading framework is the blueprint that has specified the content and guided the development of each NAEP reading assessment administered since 1992. The framework resulted from a national process involving many organizations concerned with reading education. This cooperative effort was directed by the National Assessment Governing Board (NAGB) and managed by the

Council of Chief State School Officers (CCSSO). In 2002, the NAEP reading framework was updated to provide more explicit detail regarding the assessment design.<sup>1</sup> At that time, NAGB altered slightly some of the terms used to describe elements of the reading assessment. The following description of the reading framework incorporates these changes. It should be noted, however, that this updating of the framework does not represent a change in the content or design of the NAEP reading assessment.

The framework is founded on research from the field of education that defines reading as an interactive and dynamic process involving the reader, the text, and the context of the reading experience. Reading involves the development of an understanding of text, thinking about text in different ways, and using a variety of text types for different purposes. For example, readers may read stories to enjoy and appreciate the human experience, study science texts to form new hypotheses about knowledge, or use directions to learn how to do something.

Recognizing that readers vary their approach to reading according to the demands of any particular text, the framework specifies the assessment of reading in three “contexts for reading”: reading for literary experience, reading to gain information, and reading to perform a task. Each context for reading is associated with a range of different types of texts that are included in the NAEP reading assessment. All three contexts for reading are assessed at grade 8, but reading to perform a task is not assessed at grade 4. The three contexts for reading as specified in the framework are described in figure 1.1.

---

<sup>1</sup> National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

**Figure 1.1** Descriptions of the three contexts for reading in the NAEP reading assessment

<b>Contexts for Reading</b>	
<b>Reading for literary experience</b>	<p><i>Involves the reader in exploring themes, events, characters, settings, plots, actions, and the language of literary works.</i></p> <p>Various types of texts are associated with reading for literary experience, including novels, short stories, poems, plays, legends, biographies, myths, and folktales.</p>
<b>Reading for information</b>	<p><i>Involves the engagement of the reader with aspects of the real world.</i></p> <p>Reading for information is most commonly associated with textbooks, primary and secondary sources, newspapers and magazine articles, essays, and speeches.</p>
<b>Reading to perform a task</b>	<p><i>Involves reading in order to accomplish or do something.</i></p> <p>Practical text read to perform a task may include charts, bus or train schedules, directions for games or repairs, classroom or library procedures, tax or insurance forms, recipes, voter registration materials, maps, referenda, consumer warranties, or office memos.</p>

SOURCE: National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

As readers attempt to develop understanding of text, they focus on general topics or themes, interpret and integrate ideas, make connections to background knowledge and experiences, and examine the content and structure of the text. The framework accounts for these different approaches to understanding text by

specifying four “aspects of reading” that represent the types of comprehension questions asked of students. All four aspects of reading are assessed at both grades 4 and 8 within each context of reading described above. The four aspects of reading as specified in the framework are described in figure 1.2.

**Figure 1.2 Descriptions of the four aspects of reading in the NAEP reading assessment**

Aspects of Reading	
<b>Forming a general understanding<sup>1</sup></b>	<p><i>To form a general understanding, the reader must consider the text as a whole and provide a global understanding of it.</i></p> <p>Students may be asked, for example, to demonstrate a general understanding by giving the topic of a passage, explaining the purpose of an article, or reflecting on the theme of a story.</p>
<b>Developing interpretation</b>	<p><i>To develop an interpretation, the reader must extend initial impressions to develop a more complete understanding of what was read.</i></p> <p>This process involves linking information across parts of a text as well as focusing on specific information. Questions that assess this aspect of reading include drawing inferences about the relationship of two pieces of information and providing evidence to determine the reason for an action.</p>
<b>Making reader/text connections<sup>2</sup></b>	<p><i>To make reader/text connections, the reader must connect information in the text with knowledge and experience.</i></p> <p>This process might include applying ideas in the text to the real world. All student responses to these types of questions must be text-based to receive full-credit.</p>
<b>Examining content and structure<sup>3</sup></b>	<p><i>Examining text content and structure requires critically evaluating, comparing and contrasting, and understanding the effect of such features as irony, humor, and organization.</i></p> <p>Questions used to assess this aspect of reading require readers to stand apart from the text, consider it objectively, and evaluate its quality and appropriateness. Questions ask readers to determine the usefulness of a text for a specific purpose, evaluate the language and textual elements, and think about the author's purpose and style.</p>

<sup>1</sup>This aspect of reading was referred to as "forming an initial understanding" in previous versions of the NAEP reading framework.

<sup>2</sup>This aspect of reading was referred to as "personal reflection and response" in previous versions of the NAEP reading framework.

<sup>3</sup>This aspect of reading was referred to as "demonstrating a critical stance" in previous versions of the NAEP reading framework.

SOURCE: National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

## The 2003 NAEP Reading Assessment Instrument

The NAEP reading assessment is the only federally authorized, ongoing, nationwide assessment of student reading achievement. It is governed by the framework and reflects expert perspectives on the measurement of reading comprehension. During the development process, the assessment undergoes stringent review by teachers and teacher educators, as well as by state officials and measurement specialists. All components of the assessment are evaluated for curricular relevance, developmental appropriateness, and fairness concerns.

The NAEP reading assessment measures understanding by having students read passages and answer comprehension questions. The reading passages used in the NAEP assessment are drawn from the types of books and publications that students might encounter in school, in the library, or at home. NAEP assessment developers strive to replicate authentic reading experiences in the assessment items presented to student participants. The passages students are asked to read are neither abridged nor contrived especially for the assessment. Instead, full-length reading selections are reprinted in test booklets to resemble as closely as possible the format of their

original publication. To demonstrate their comprehension of these passages, students answer a combination of multiple-choice and constructed-response questions. The multiple-choice questions include four options from which students are asked to select the best answer. The constructed-response questions require students to write their own responses. Short constructed-response questions can be completed in no more than a few sentences, while extended constructed-response questions may require students to provide responses as long as a paragraph or a full page. Both types of constructed-response questions require students to support their answers by using information in the reading passage.

In order to ensure reliable and valid scoring of constructed-response questions, a unique scoring guide, describing the specific criteria for assigning a score level to each student's response, is developed for each question. Expert scorers go through extensive training to understand how to apply these scoring criteria fairly and consistently. Scorers are consistently monitored to ensure that scoring standards are being applied appropriately and to ensure a high degree of scorer agreement (i.e., interrater reliability). In addition, for those constructed-response questions that were used in previous assessments, monitoring of scorers includes checking to make sure that scoring standards remain consistent from year to year.

At each grade, the entire reading assessment is divided into sections referred to as blocks. Each block contains at least one text and a related set of approximately 10 to 12 comprehension

questions (a combination of multiple-choice and constructed-response). Most of the blocks are presented to students as 25-minute timed sections, but some are presented as 50-minute timed sections. The total number of blocks that make up the NAEP reading assessment at each grade are as follows:

Grade 4—five 25-minute literary blocks and five 25-minute informative blocks

Grade 8—four 25-minute literary blocks, four 25-minute informative blocks, four 25-minute task blocks, and one 50-minute informative block

In order to minimize the burden on any individual student, NAEP uses a procedure referred to as matrix sampling, in which an individual student is administered only a small portion of the entire assessment at any grade. For example, at grade 4, students are given a test booklet that contains only two 25-minute blocks. At grade 8, students are given a test booklet that contains either two 25-minute blocks or one 50-minute block. Because each block is administered to a representative sample at each grade, the results can then be combined to produce average group and subgroup results based on the entire assessment. In addition to the two 25-minute blocks or one 50-minute block in each student's test booklet, students are asked to complete two sections of background questions that ask about their background and home or school experiences related to reading achievement. The time required for each student to participate in the NAEP reading assessment is approximately one hour.



## Description of School and Student Samples

The NAEP 2003 reading assessment was administered to fourth- and eighth-graders at the national and the state levels. At the national level, results are reported for both public and nonpublic school students. At the state or jurisdiction level, results are reported only for public school students. In order to obtain a representative sample of students for reporting national and state or jurisdiction results, approximately 188,000 fourth-graders from 7,500 schools and 155,000 eighth-graders from 6,100 schools were sampled and assessed. All 50 states and 3 jurisdictions participated and met the minimum guidelines for reporting their results in 2003. The national samples were larger in 2002 and 2003 than in previous assessment years because they were based on the combined sample of public school students assessed in each participating state, plus an additional sample from nonpublic schools. In 1992–2000 the national samples were drawn separately from the state samples and were smaller than the samples resulting from aggregating the state samples. Each selected school that participated in the assessment and each student assessed represents a portion of the population of interest. For information on sample sizes and participation rates for the nation and by state or jurisdiction, see tables A.6–A.9 in appendix A.

Results from the 2002 and 2003 Trial Urban District Assessment (TUDA) are reported for the participating districts for public school students at grades 4 and 8. The TUDA employed larger-than-usual samples within the districts, making reliable district-level data possible. The samples were also large enough to provide reliable estimates on subgroups

within the districts, such as female students or Hispanic students.

## Reporting the Assessment Results

Results from the NAEP reading assessment are presented in terms of scale scores and percentages of students attaining achievement levels. The scale score results, indicating how much students *know and can do* in reading, are presented as average scale scores and as scale scores at selected percentiles. The achievement-level results indicate the degree to which student performance meets the standards set for what they *should know and be able to do*. Results are reported only for groups or subgroups of students; individual student performance cannot be reported based on the NAEP assessment.

Average scale score results are based on the NAEP reading scale, which ranges from 0 to 500. In order to calculate students' average scores on the NAEP reading assessment, the analysis begins by determining the percentages of students responding correctly to each multiple-choice question and the percentages of students responding at each score level for each constructed-response question. The analysis entails summarizing the results on separate subscales for each reading context (reading for literary experience, reading for information, and reading to perform a task) and then combining the separate scales to form a single composite reading scale. The relative contribution of each reading purpose at each grade is displayed in table I.1. (See appendix A for more information on scaling procedures.)

Achievement-level results are presented in terms of reading achievement levels as authorized by the NAEP legislation and adopted by NAGB. For each

**Table 1.1** Percentage weighting of the “context for reading” subscales on the NAEP composite reading scale, grades 4 and 8

NAEP Reading Subscales	Reading for literary experience	Reading for information	Reading to perform a task
Grade 4	55	45	—
Grade 8	40	40	20

— Not available. Not assessed at grade 4.

SOURCE: National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

grade assessed, NAGB has adopted three achievement levels: *Basic*, *Proficient*, and *Advanced*. For reporting purposes, achievement-level cut scores are placed on the reading scale, resulting in four ranges: below *Basic*, *Basic*, *Proficient*, and *Advanced*. The achievement-level results are then reported as percentages of students within each achievement-level range, as well as the percentage of students at or above *Basic* and at or above *Proficient*.

### The Setting of Achievement Levels

The 1988 NAEP legislation that created NAGB directed the Board to identify “appropriate achievement goals . . . for each subject area” that NAEP measures.<sup>2</sup> The NAEP 2001 reauthorization reaffirmed many of the Board’s statutory responsibilities, including developing “appropriate student achievement levels for each grade or age in each subject area to be tested. . . .”<sup>3</sup> In order to follow this directive and achieve the mandate of the 1988 statute “to improve the form and use of NAEP results,” NAGB undertook the development of student perfor-

mance standards (called “achievement levels”).<sup>4</sup> Since 1990, the Board has adopted achievement levels in mathematics, reading, U.S. history, world geography, science, writing, and civics.

The Board defined three levels for each grade: *Basic*, *Proficient*, and *Advanced*. The *Basic* level denotes partial mastery of the knowledge and skills that are fundamental for proficient work at a given grade. The *Proficient* level represents solid academic performance. Students reaching this level demonstrate competency over challenging subject matter. The *Advanced* level presumes mastery of both the *Basic* and *Proficient* levels and represents superior performance. Figure 1.3 presents the policy definitions of the achievement levels that apply across grades and subject areas. The policy definitions guided the development of the reading achievement levels, as well as the achievement levels established in all other subject areas assessed by NAEP. Adopting three levels of achievement for each grade signals the importance of looking at more than one standard of

<sup>2</sup> National Assessment of Educational Progress Improvement Act, Pub. L. No. 100–297, 20 U.S.C. § 1221 *et seq.* (1988).

<sup>3</sup> No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425 (2002).

<sup>4</sup> National Assessment of Educational Progress Improvement Act, Pub. L. No. 100–297, 20 U.S.C. § 1221 *et seq.* (1988).

**Figure 1.3 Policy definitions of the three NAEP achievement levels**

Achievement Levels	
<b>Basic</b>	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.
<b>Proficient</b>	This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
<b>Advanced</b>	This level signifies superior performance.

SOURCE: National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

performance. In the Board's view, the overall achievement goal for students is performance that qualifies at the *Proficient* level or higher as measured by NAEP. The *Basic* level is not the desired goal, but represents partial mastery that is a step toward *Proficient*.

The achievement levels in this report were adopted by the Board based on a standard-setting process designed and conducted under a contract with ACT. To develop these levels, ACT convened a cross section of educators and interested citizens from across the nation and asked them to judge what students should know and be able to do relative to a body of content reflected in the reading framework. This achievement-level-setting process was reviewed by numerous individuals including policymakers, representatives of professional organizations, teachers, parents, and other members of the general public. Prior to adopting these levels of student achievement, NAGB engaged a large number of persons to comment on the recommended levels and to review the results.

The results of the achievement-level-setting process, after NAGB's approval, became a set of achievement-level descriptions and a set of achievement-level cut scores. The cut scores are the scores on the 0–500 NAEP reading scale that define the lower boundaries of *Basic*, *Proficient*, and *Advanced* performance levels at grades 4, 8, and 12.

### **Reading Achievement-Level Descriptions for Each Grade**

Specific definitions of the *Basic*, *Proficient*, and *Advanced* reading achievement levels for grades 4 and 8 are presented in figures 1.4 and 1.5. The achievement levels are cumulative; therefore, students performing at the *Proficient* level also display the competencies associated with the *Basic* level, and students at the *Advanced* level also demonstrate the competencies associated with both the *Basic* and the *Proficient* levels. For each achievement level listed in figures 1.4 and 1.5, the scale score that corresponds to the lowest score within that level on the NAEP reading scale is shown in parentheses. For example, in figure 1.4 the scale score of 238 corresponds to the lowest score in the range defining the grade 4 *Proficient* level of achievement in reading.

**Figure 1.4** Descriptions of NAEP reading achievement levels, grade 4

<b>Grade 4 Achievement Levels</b>	
<b>Basic</b> (208)	<p>Fourth-grade students performing at the <i>Basic</i> level should demonstrate an understanding of the overall meaning of what they read. When reading text appropriate for fourth graders, they should be able to make relatively obvious connections between the text and their own experiences, and extend the ideas in the text by making simple inferences.</p> <p>For example, when reading <b>literary</b> text, they should be able to tell what the story is generally about—providing details to support their understanding—and be able to connect aspects of the stories to their own experiences.</p> <p>When reading <b>informational</b> text, <i>Basic</i>-level fourth graders should be able to tell what the selection is generally about or identify the purpose for reading it, provide details to support their understanding, and connect ideas from the text to their background knowledge and experiences.</p>
<b>Proficient</b> (238)	<p>Fourth-grade students performing at the <i>Proficient</i> level should be able to demonstrate an overall understanding of the text, providing inferential as well as literal information. When reading text appropriate to fourth grade, they should be able to extend the ideas in the text by making inferences, drawing conclusions, and making connections to their own experiences. The connections between the text and what the student infers should be clear.</p> <p>For example, when reading <b>literary</b> text, <i>Proficient</i>-level fourth graders should be able to summarize the story, draw conclusions about the characters or plot, and recognize relationships such as cause and effect.</p> <p>When reading <b>informational</b> text, <i>Proficient</i>-level students should be able to summarize the information and identify the author's intent or purpose. They should be able to draw reasonable conclusions from the text, recognize relationships such as cause and effect or similarities and differences, and identify the meaning of the selection's key concepts.</p>
<b>Advanced</b> (268)	<p>Fourth-grade students performing at the <i>Advanced</i> level should be able to generalize about topics in the reading selection and demonstrate an awareness of how authors compose and use literary devices. When reading text appropriate to fourth grade, they should be able to judge texts critically and, in general, give thorough answers that indicate careful thought.</p> <p>For example, when reading <b>literary</b> text, <i>Advanced</i>-level students should be able to make generalizations about the point of the story and extend its meaning by integrating personal experiences and other readings with ideas suggested by the text. They should be able to identify literary devices such as figurative language.</p> <p>When reading <b>informational</b> text, <i>Advanced</i>-level fourth graders should be able to explain the author's intent by using supporting material from the text. They should be able to make critical judgments of the form and content of the text and explain their judgments clearly.</p>

SOURCE: National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

**Figure 1.5 Descriptions of NAEP reading achievement levels, grade 8**

<b>Grade 8 Achievement Levels</b>	
<b>Basic</b> (243)	<p>Eighth-grade students performing at the <i>Basic</i> level should demonstrate a literal understanding of what they read and be able to make some interpretations. When reading text appropriate to eighth grade, they should be able to identify specific aspects of the text that reflect the overall meaning, extend the ideas in the text by making simple inferences, recognize and relate interpretations and connections among ideas in the text to personal experience, and draw conclusions based on the text.</p> <p>For example, when reading <b>literary</b> text, <i>Basic</i>-level eighth graders should be able to identify themes and make inferences and logical predictions about aspects such as plot and characters.</p> <p>When reading <b>informational</b> text, they should be able to identify the main idea and the author's purpose. They should make inferences and draw conclusions supported by information in the text. They should recognize the relationships among the facts, ideas, events, and concepts of the text (e.g., cause and effect, order).</p> <p>When reading <b>practical</b> text, they should be able to identify the main purpose and make predictions about the relatively obvious outcomes of procedures in the text.</p>
<b>Proficient</b> (281)	<p>Eighth-grade students performing at the <i>Proficient</i> level should be able to show an overall understanding of the text, including inferential as well as literal information. When reading text appropriate to eighth grade, they should be able to extend the ideas in the text by making clear inferences from it, by drawing conclusions, and by making connections to their own experiences—including other reading experiences. <i>Proficient</i> eighth graders should be able to identify some of the devices authors use in composing text.</p> <p>For example, when reading <b>literary</b> text, students at the <i>Proficient</i> level should be able to give details and examples to support themes that they identify. They should be able to use implied as well as explicit information in articulating themes; to interpret the actions, behaviors, and motives of characters; and to identify the use of literary devices such as personification and foreshadowing.</p> <p>When reading <b>informational</b> text, they should be able to summarize the text using explicit and implied information and support conclusions with inferences based on the text.</p> <p>When reading <b>practical</b> text, <i>Proficient</i>-level students should be able to describe its purpose and support their views with examples and details. They should be able to judge the importance of certain steps and procedures.</p>
<b>Advanced</b> (323)	<p>Eighth-grade students performing at the <i>Advanced</i> level should be able to describe the more abstract themes and ideas of the overall text. When reading text appropriate to eighth grade, they should be able to analyze both meaning and form and support their analyses explicitly with examples from the text, and they should be able to extend text information by relating it to their experiences and to world events. At this level, student responses should be thorough, thoughtful, and extensive.</p> <p>For example, when reading <b>literary</b> text, <i>Advanced</i>-level eighth graders should be able to make complex, abstract summaries and theme statements. They should be able to describe the interactions of various literary elements (i.e., setting, plot, characters, and theme) and explain how the use of literary devices affects both the meaning of the text and their response to the author's style. They should be able to critically analyze and evaluate the composition of the text.</p> <p>When reading <b>informational</b> text, they should be able to analyze the author's purpose and point of view. They should be able to use cultural and historical background information to develop perspectives on the text and be able to apply text information to broad issues and world situations.</p> <p>When reading <b>practical</b> text, <i>Advanced</i>-level students should be able to synthesize information that will guide their performance, apply text information to new situations, and critique the usefulness of the form and content.</p>

SOURCE: National Assessment Governing Board. (2002). *Reading Framework for the 2003 National Assessment of Educational Progress*. Washington, DC: Author.

## Trial Status of Achievement Levels

The law requires that the achievement levels are to be used on a trial basis until the Commissioner of Education Statistics determines “that such levels are reasonable, valid, and informative to the public.”<sup>5</sup> Until that determination is made, the law requires the Commissioner and the Board to state clearly the trial status of the achievement levels in all NAEP reports. In 1993, the first of several congressionally mandated evaluations of the achievement-level-setting process concluded that the procedures used to set the achievement levels were flawed and that the percentage of students at or above any particular achievement-level cut point may be underestimated.<sup>6</sup> Others have critiqued these evaluations, asserting that the weight of the empirical evidence does not support such conclusions.<sup>7</sup>

In response to the evaluations and critiques, NAGB sponsored an additional study of the 1992 reading achievement levels before deciding to use them for reporting NAEP 1994 results.<sup>8</sup> When reviewing the findings of this study, the

National Academy of Education (NAE) panel expressed concern about what it saw as a “confirmatory bias” in the study and about the inability of this study to “address the panel’s perception that the levels had been set too high.”<sup>9</sup> In 1997, the NAE panel summarized its concerns with interpreting NAEP results based on the achievement levels as follows:

First, the potential instability of the levels may interfere with the accurate portrayal of trends. Second, the perception that few American students are attaining the higher standards we have set for them may deflect attention to the wrong aspects of education reform. The public has indicated its interest in benchmarking against international standards, yet it is noteworthy that when American students performed very well on a 1991 international reading assessment, these results were discounted because they were contradicted by poor performance against the possibly flawed NAEP reading achievement levels in the following year.<sup>10</sup>

<sup>5</sup> No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

<sup>6</sup> United States General Accounting Office. (1993). *Education Achievement Standards: NAGB’s Approach Yields Misleading Interpretations*. U.S. General Accounting Office Report to Congressional Requestors. Washington, DC: Author.

National Academy of Education. (1993). *Setting Performance Standards for Achievement: A Report of the National Academy of Education Panel on the Evaluations of the NAEP Trial State Assessment: An Evaluation of the 1992 Achievement Levels*. Stanford, CA: Author.

<sup>7</sup> Cizek, G. (1993). *Reactions to National Academy of Education Report*. Washington, DC: National Assessment Governing Board.

Kane, M. (1993). *Comments on the NAE Evaluation of the NAGB Achievement Levels*. Washington, DC: National Assessment Governing Board.

<sup>8</sup> American College Testing. (1995). *NAEP Reading Revisited: An Evaluation of the 1992 Achievement Level Descriptions*. Washington, DC: National Assessment Governing Board.

<sup>9</sup> National Academy of Education. (1996). Reading Achievement Levels. In *Quality and Utility: The 1994 Trial State Assessment in Reading. The Fourth Report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment*. Stanford, CA: Author.

<sup>10</sup> National Academy of Education. (1997). *Assessment in Transition: Monitoring the Nation’s Educational Progress*, p. 99. Mountain View, CA: Author.

NCES and NAGB have sought and continue to seek new and better ways to set performance standards for NAEP.<sup>11</sup> For example, NCES and NAGB jointly sponsored a national conference that explored many issues related to standard setting in large-scale assessments.<sup>12</sup> Although new directions were presented and discussed, a proven alternative to the current process has not yet been identified. NCES and NAGB continue to call on the research community to assist in finding ways to improve standard setting for reporting NAEP results.

The most recent congressionally mandated evaluation conducted by the National Academy of Sciences (NAS) relied on prior studies of achievement levels, rather than carrying out new evaluations, on the grounds that the process has not changed substantially since the initial problems were identified. Instead, the NAS panel studied the development of the 1996 science achievement levels. The NAS panel basically concurred with earlier congressionally mandated studies. The panel concluded that “NAEP’s current achievement-level-setting procedures remain fundamentally flawed. The judgment tasks are difficult

and confusing; raters’ judgments of different item types are internally inconsistent; appropriate validity evidence for the cut scores is lacking; and the process has produced unreasonable results.”<sup>13</sup>

The NAS panel accepted the continuing use of achievement levels in reporting NAEP results on a trial basis, until such time as better procedures can be developed. Specifically, the NAS panel concluded that “. . . tracking changes in the percentages of students performing at or above those cut scores (or in fact, any selected cut scores) can be of use in describing changes in student performance over time.”<sup>14</sup>

NAGB urges all who are concerned about student performance levels to recognize that the use of these achievement levels is a developing process and is subject to various interpretations. NAGB and NCES believe that the achievement levels are useful for reporting trends in the educational achievement of students in the United States.<sup>15</sup> In fact, achievement-level results have been used in reports by the President of the United States, the Secretary of Education, state governors, legislators, and members of Congress. Government leaders in the

---

<sup>11</sup> Reckase, M. D. (2000). *The Evolution of the NAEP Achievement Levels Setting Process: A Summary of the Research and Development Efforts Conducted by ACT*. Iowa City, IA: ACT, Inc.

<sup>12</sup> National Assessment Governing Board and National Center for Education Statistics. (1995). *Proceedings of the Joint Conference on Standard Setting for Large-Scale Assessments of the National Assessment Governing Board (NAGB) and the National Center for Education Statistics (NCES)*. Washington, DC: Government Printing Office.

<sup>13</sup> Pellegrino, J. W., Jones, L. R., and Mitchell, K. J. (Eds.). (1998). *Grading the Nation’s Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress*. Committee on the Evaluation of National Assessments of Educational Progress, Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

<sup>14</sup> *Ibid.*, 176.

<sup>15</sup> Forsyth, R. A. (2000). A Description of the Standard-Setting Procedures Used by Three Standardized Test Publishers. In *Student Performance Standards on the National Assessment of Educational Progress: Affirmations and Improvements*. Washington, DC: National Assessment Governing Board.

Nellhaus, J. M. (2000). States with NAEP-Like Performance Standards. In *Student Performance Standards on the National Assessment of Educational Progress: Affirmations and Improvements*. Washington, DC: National Assessment Governing Board.

nation and in more than 40 states use these results in their annual reports. However, based on the congressionally mandated evaluations so far, NCES agrees with the NAS panel's recommendation that caution needs to be exercised in the use of the current achievement levels. NCES has concluded that these achievement levels should continue to be used on a trial basis and be interpreted with caution.

### **Interpreting NAEP Results**

The average scores and percentages presented in this report are estimates based on samples of students rather than on entire populations. Moreover, the collection of questions used at each grade level is but a sample of the many questions that could have been asked to assess the skills and abilities described in the NAEP reading framework. As such, the results are subject to a measure of uncertainty, reflected in the standard error of the estimates—a range of up to a few points above or below the score or percentage—which accounts for potential score or percentage fluctuation due to sampling and measurement error. The estimated standard errors for the estimated scale scores and percentages in this report are accessible through the NAEP Data Tool on the NAEP web site (<http://nces.ed.gov/nationsreportcard/naepdata/>). Examples of these estimated standard errors are also provided in appendix A of this report.

The differences between scale scores and between percentages discussed in the following chapters take into account the standard errors associated with the estimates. Comparisons are based on statistical tests that consider both the

magnitude of the difference between the group average scores or percentages and the standard errors of those statistics. Estimates based on smaller subgroups are likely to have relatively large standard errors. As a consequence, some seemingly large differences may not be statistically significant. That is, it cannot be determined whether these differences are due to the particular make-up of the samples of students who were selected, or to true differences in the population of interest. When this is the case, the term “apparent difference” or “no measurable difference” is used in this report. Differences between scores or between percentages are discussed in this report only when they are significant from a statistical perspective.

Beginning in 2002, the NAEP national sample was obtained by aggregating the samples from each state, rather than obtaining an independently selected national sample. Consequently, the national sample size increased and smaller differences between years or between subgroups of students were found to be statistically significant than would have been detected in previous assessment years. In keeping with past practice, all statistically significant differences are indicated in this report. All differences reported are significant at the .05 level with appropriate adjustments for multiple comparisons. 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 dependable differences in average scores or percentages to help inform dialogue among policymakers, educators, and the public.



While the score ranges at each grade in reading are identical, the scale was derived independently at each grade. Therefore, average scale scores across grades cannot be compared. For example, equal scale scores on the grade 4 and grade 8 scales do not imply equal levels of reading achievement.

Comparisons of performance results may be affected by changes in exclusion rates for students with disabilities and limited-English-proficient students in the NAEP samples. Percentages of students excluded from NAEP may vary considerably across states or districts, as well as across years. Comparisons of achievement results should be interpreted with caution if the exclusion rates vary widely. Percentages of students with disabilities and limited-English-proficient students identified, excluded, and assessed are presented in tables A.10–A.21 in appendix A.

The results presented are meant to describe some aspects of the condition of education. They are best viewed as suggesting various ideas to be further examined in light of other data, including state and local data, and in the context of the large research literature elaborating on the many factors contributing to educational achievement.

However, some readers are tempted to make unwarranted causal inferences from simple cross tabulations. At the risk of sounding dogmatic, it is almost never the case that a simple cross tabulation of any variable with a measure of educational achievement is conclusive proof that differences in that variable are a cause of differential educational achievement. The old adage that “correlation is not causation” is a wise precaution to be

kept in mind when viewing the results presented here. Experienced researchers routinely formulate multiple hypotheses to take these possibilities into account and readers of this volume are encouraged to do likewise.

Additional NAEP data are available in the NAEP data tool and in restricted-access research databases. Researchers and policy analysts are free to make use of the data (subject to various confidentiality restrictions) as they wish. However, as part of the Institute for Education Sciences, NCES has a responsibility to try to discourage misleading inferences from the data presented and to educate the public on the difficulty of making valid causal inferences in a field as complex as education.

### **Overview of the Remaining Report**

This report describes the reading performance of fourth- and eighth-graders in the nation, in participating states and other jurisdictions, in large central city school districts, and in selected urban school districts. Chapter 2 presents overall reading scale-score and achievement-level results across years for both the nation and participating states and other jurisdictions. Chapter 3 discusses national results for subgroups of students by gender, race/ethnicity, eligibility for free/reduced-price school lunch, parents' highest level of education (for grade 8 only), type of school (public and nonpublic), and school's type of location (central city, urban fringe/large town, rural/small town). State and jurisdiction results are reported by gender, race/ethnicity, and eligibility for free/reduced-price school lunch. Overall and subgroup results for selected urban districts are presented in chapter 4.

Chapter 5 presents sample assessment questions and student responses at each grade level, including samples of multiple-choice and constructed-response questions. A table showing the percentage of students who answered the question successfully accompanies each sample question. In addition, item maps for each grade level describe the skill or ability needed to answer particular reading questions and show the score points at which individual students had a high probability of successfully answering particular questions, thereby indicating the relative difficulty of each question.

The appendices of this report contain information to expand the results presented in chapters 2–5. Appendix A contains an overview of assessment development, sampling, administration, and analysis procedures. Appendix B presents the percentages of students in each of the subgroups reported for the nation, states and other jurisdictions, and districts. Appendix C includes additional state-level results by subgroup. Appendix D shows state-level and district-level contextual data from sources other than NAEP. Appendix E contains the reading passages corresponding with the sample questions discussed in chapter 5.