DOCUMENT RESUME

ED 398 139 SO 026 695

AUTHOR Beatty, Alexandra S.; And Others
TITLE NAEP 1994 U.S. History Report Card. Findings from the National Assessment of Educational Progress.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
REPORT NO ISBN-0-16-048645-9; NCES-96-085
PUB DATE Apr 96
NOTE 134p.; For related documents, see ED 389 666 and SO 026 762.
PUB TYPE Information Analyses (070) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Academic Achievement; Curriculum Based Assessment; *Educational Assessment; Elementary Secondary Education; Evaluation Methods; Grade 4; Grade 8; Grade 12; History Instruction; *National Programs; Standards; *Student Evaluation; Tests; *United States History
IDENTIFIERS *National Assessment of Educational Progress

ABSTRACT

This book describes results from the 1994 National Assessment of Educational Progress (NAEP) assessment in U.S. history, conducted at grades 4, 8, and 12. Included in this report card are the results of students' achievement at each grade and within various subgroups of the general population. The report discusses the relationships between student performance and instructional and home background variables. This information gives educators a context for evaluating the U.S. history achievement of students and the results that may be used to guide reform efforts. Chapters include: (1) "NAEP 1994 U.S. History Assessment"; (2) "U.S. History Results for the Nation and Regions"; (3) "U.S. History Achievement Levels"; (4) "Contexts in which Students Learn History"; and (5) "What Students Know and Can Do in U.S. History." A conclusion, three appendices, 52 tables, and 13 figures complete the book. (EH)

NATIONAL CENTER FOR EDUCATION STATISTICS

NAEP 1994 U.S. History Report Card

FINDINGS FROM THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

U.S. DEPARTMENT OF EDUCATION
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT

What is The Nation's Report Card?

THE NATION'S REPORT CARD, the National Assessment of Educational Progress (NAEP), is the only 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, 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. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

In 1988, Congress established the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The Board is responsible for selecting the subject areas to be assessed from among those included in the National Education Goals; for setting appropriate student performance levels; for developing assessment objectives and test specifications through a national consensus approach; for designing the assessment methodology; for developing guidelines for reporting and disseminating NAEP results; for developing standards and procedures for interstate, regional, and national comparisons; for determining the appropriateness of test items and ensuring they are free from bias; and for taking actions to improve the form and use of the National Assessment.

The National Assessment Governing Board

NATIONAL CENTER FOR EDUCATION STATISTICS

NAEP 1994 U.S. History Report Card

Findings from the National Assessment of Educational Progress

Alexandra S. Beatty
Clyde M. Reese
Hilary R. Persky
Peggy Carr

April 1996

Office of Educational Research and Improvement
U.S. Department of Education

Prepared by Educational Testing Service under contract with the National Center for Education Statistics

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Jeanne E. Griffith
Acting Commissioner

Education Assessment Division
Gary W. Phillips
Associate Commissioner

April 1996

FOR MORE INFORMATION:

For ordering information on this report, write:

National Library of Education
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, D.C. 20208-5641

or call 1-800-424-1616 (in the Washington, DC, metropolitan area
call 202-219-1651).

The work upon which this publication is based was performed for the
National Center for Education Statistics, Office of Educational Research
and Improvement, by Educational Testing Service.

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-048645-9

TABLE OF CONTENTS

APPENDICES

TABLES

FIGURES

EXECUTIVE SUMMARY

If policymakers, educators, and concerned citizens are to improve the United States educational system, they need valid and reliable information about the strengths and weaknesses of American students and the instructional factors that are related to differing levels of performance. For over twenty-five years, the National Assessment of Educational Progress (NAEP) has provided such information. NAEP assessments have probed student abilities in a variety of subject areas, reporting both on what students know and can do and on the relationships between instructional, institutional, and background variables and differing levels of educational achievement. As the nation's foremost ongoing education survey, NAEP results track trends in student performance and allow concerned readers to evaluate whether America's children are developing the skills and knowledge essential for effective participation in the economy and the polity.

In 1994, NAEP conducted national assessments in reading, geography, and United States history at grades 4, 8, and 12. The United States history results included in this Report Card describe students' achievement at each grade and within various subgroups of the general population. In addition, the report discusses the relationships between student performance and instructional and home background variables. Taken together, this information gives educators a context for evaluating the U.S. history achievement of students, and results that may be used to guide reform efforts.

Student performance is summarized on the NAEP U.S. history scale, which ranges from 0 to 500. In addition, results for each grade are reported according to three achievement levels: *Basic*, *Proficient*, and *Advanced*. These achievement levels are based on collective judgments about what students should know and be able to do in U.S. history. The *Basic* level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade. The *Proficient* level represents solid academic performance and demonstrated competence over challenging subject matter. The *Advanced* level signifies superior performance.

Major Findings for the Nation

- The *Proficient* achievement level — defined as signifying solid academic performance and demonstrated competence over challenging subject matter — was reached by only 17 percent of fourth graders, 14 percent of eighth graders, and 11 percent of twelfth graders.

- Fewer than half the grade 12 students in the assessment were able to reach the *Basic* level. At grades 4 and 8, over 60 percent of assessed students demonstrated this level of performance.

- On individual assessment tasks, students demonstrated a range of competencies. For example:

  - At grade 4, 87 percent of assessed students identified Martin Luther King Jr's "I Have A Dream" speech; 45 percent identified the primary purpose of the Bill of Rights; 32 percent knew that New York was one of the colonies that took part in the American Revolution; and 11 percent described some ways in which women's roles in the workforce have changed over the past 100 years.

  - At grade 8, 80 percent of the students knew that the song "O Freedom" was used by people in the Civil Rights movement; 71 percent identified Thomas Jefferson as the author of the Declaration of Independence; 41 percent associated the Lend-Lease Act, the Yalta Conference, and the dropping of the atomic bomb with the Second World War; and 10 percent wrote a description of the debate at the Constitutional Convention that led to the Connecticut Compromise.

  - At grade 12, 74 percent of the students determined the impact of the launch of the Soviet Sputnik satellite on United States politics; 41 percent identified the purpose of the Monroe Doctrine; 29 percent explained the effect of an economic or technological change on the nature of farming in America; and 22 percent successfully compared Franklin Roosevelt's 1933 and 1937 inauguration speeches.

Major Findings for Student Subgroups

- As has been the case in other NAEP assessments, there were statistically significant differences in the performance of major subgroups of the population. For example, at all grades White and Asian students had higher scores than did their Black and Hispanic counterparts.

- Consistent with findings in other assessments, there was a strong relationship between differing levels of parental education and performance on the NAEP U.S. history assessment. As a general rule, the more education students' parents had received, the better the students performed on the assessment.

- On the overall U.S. history scale, there were no statistically significant performance differences between male and female students at grades 4 and 8. At grade 12, male students performed at a higher level than females. However, gender differences were not consistent across areas of U.S. history. At all grades, males outperformed females on tasks assessing the historical theme of "The Changing Role of America in the World." On the other hand, eighth- and twelfth-grade females performed better than their male counterparts on the "Gathering and Interaction of Peoples, Cultures, and Ideas" theme.

- At all three grades, students attending nonpublic schools performed at a higher level than did students attending public schools.

Contextual Factors Related to United States History Performance

A diverse range of home and school factors influence the ways and extent to which students learn history. Students who participated in the NAEP assessment were asked to complete questionnaires about their home and school experiences related to history learning. Also, questionnaires about students' instructional experiences were completed by their teachers. The results of these surveys help provide a context for interpreting the assessment scores, and provide policymakers with information about variables that are positively and negatively related to history achievement.

- Over 40 percent of the students at grades 4 and 8 and 25 percent of the students at grade 12 reported watching four or more hours of television each day. In most cases, the more television students reported watching the worse they performed on the U.S. history assessment.

- Fifty-four percent of the fourth graders, 38 percent of the eighth graders, and 31 percent of the twelfth graders reported discussing their studies at home daily. By contrast, 18, 22, and 25 percent of the students at each grade respectively reported never or hardly ever discussing their studies. Students who reported no regular discussions had lower average scale scores than all other students.

- United States history instruction was limited for grade 4 students. Only 7 percent of the fourth graders assessed had teachers who reported that U.S. history is the focus of their social studies teaching. Conversely, most students in grade 8 were taking a course in U.S. history, and most twelfth graders had taken such a course in grade 11.

- History homework was also limited at grade 8. Forty-eight percent of students who were taking U.S. history reported that they did one-half hour or less of history homework each week. These students had lower scores on the assessment than did students who did one or two hours of homework each week.

- Almost half the students had teachers who reported using textbooks on a daily basis. Also, 62 percent of the students at grade 4 and 23 percent of the students at grade 8 had teachers who reported that they never or hardly ever use primary documents in their teaching.

About This Report

As the Nation's Report Card in United States history, this document provides a broad examination of history learning. In addition, specific aspects of students' performance and their experiences at home and school are reviewed in some depth. As such, this report provides a portrait of what students know and can do in history, as well as the contexts in which they have developed their history knowledge and skills.

Chapter 1 presents the overview of the NAEP 1994 U.S. history assessment — its content framework, design, and administration. Also included in Chapter 1 are example questions and student responses from the assessment. Chapter 2 provides overall average scale score results for the nation, regions, and subgroups of students. Chapter 3 describes student performance in terms of achievement levels. Chapter 4 describes contextual factors related to students' performance. Finally, Chapter 5 describes the specific abilities demonstrated by students in the NAEP 1994 U.S. history assessment and reports student performance in different thematic areas of U.S. history.

CHAPTER 1

NAEP 1994 U.S. History Assessment

Introduction

"I hope that through my teaching I can open young people's eyes about our nation's past and encourage them to always ask, 'Why is X the way it is? How did X get to be so?' If I can do that, I've helped train a generation of inquisitive, curious minds which will be able to pursue achievement in a variety of fields."[1]

Knowledge of United States history is an important component of effective citizenship. A thorough grasp of our country's struggles, successes, and failures, and the skills to interpret them, better enable young people to make informed and intelligent decisions about contemporary issues. Nourishing the curiosity children exhibit about major events, customs and institutions, and the families and individuals that comprise United States history, creates a valuable resource for our nation's future.

While there is increasingly wide agreement about the importance of historical study for students, there has been little consensus about history's place in the curriculum or about what should be taught in history classrooms. During the decades after World War II, history was frequently pushed aside to make room for social science classes and for courses designed to help students cope with issues and problems in their everyday lives. Many educators believed that history could be embedded in a wider social studies curriculum.

The inclusion of history in the *Goals 2000* list of key subjects in which American students should be expected to show mastery by the year 2000 was an important signal of the education community's renewed commitment to history's importance.[2] Recent evidence suggests that schools are making their history curricula more rigorous and that history teachers are maintaining higher standards.[3] This emphasis on high expectations — also reflected in the National Assessment of Educational Progress (NAEP) U.S. history assessments — has pushed attempts to improve history instruction to the forefront of the educational reform movement. Furthermore, the *NAEP 1994 U.S. History Framework,* the guiding document in the preparation of the NAEP assessments, represents a consensus among historians and history educators about the rigorous skills and content knowledge that all students should be able to demonstrate.

The renewed interest in history instruction underscores the need for accurate information about what students know and can do in U.S. history and about which instructional background factors are related to high levels of history achievement. The NAEP 1994 U.S. history assessment provides a detailed portrait of student performance that can serve as a benchmark for educators, parents, policymakers, and the general public.

Overview of the 1994 National Assessment of Educational Progress (NAEP)

A project of the National Center for Education Statistics (NCES), NAEP collects information about what students in the United States know and can do in various school subjects. Since its initiation by Congress in 1969, NAEP has carried out its federally supported mandate as the only ongoing national assessment of student achievement. Public and nonpublic school students in grades 4, 8, and 12 are regularly sampled and assessed in reading, mathematics, history, geography, and other subject areas. The assessments are based on frameworks that prescribe the content of the assessment. The content of the assessments attempts to maintain a balance among current instructional efforts, curriculum reform, research results, and desirable levels of achievement.

The NAEP 1994 U.S. history assessment was administered to national samples of fourth-, eighth-, and twelfth-grade students attending public and nonpublic schools. This report presents the results of this assessment. Each participant was asked to answer a set of background questions and complete a series of history exercises. Approximately 22,000 students were assessed. Students' performance is described on the NAEP U.S. history scale ranging from 0 to 500 and in relation to three achievement levels: *Basic, Proficient,* and *Advanced.* The average performance of students at each of the three grades and of specific subgroups of the student population is presented.

The NAEP 1994 U.S. History Framework

The structure and content of the U.. history assessment were guided by the *NAEP 1994 U.S. History Framework.*[4] Although U.S. history was assessed by NAEP in 1988, a rigorous new framework was developed for the 1994 assessment. The development of the new framework, as well as the detailed specifications for the assessment, were managed by the Council of Chief State School Officers (CCSSO) under the direction of the National Assessment Governing Board (NAGB). Approximately fifty professional historians, educators, administrators, and other interested individuals worked to achieve consensus on the general goals as well as the specific language for the framework document. In addition, several hundred educational experts and interested members of the public contributed to the process, either by participating in public hearings or by reviewing drafts of the documents.

The historians and educators involved in the consensus process believed that an entirely new framework was needed to address the range of historical content and skills that students should possess. The resulting framework, which called for the assessment of a broad range of outcomes, represented an ambitious vision both of what students should know and be able to do in U.S. history, and of the ways in which those competencies should be tested.

The 1994 framework is organized around three concepts or dimensions: major themes of U.S. history, chronological periods of U.S. history, and ways of knowing and thinking about U.S. history. These organizing concepts grew out of the goals for the assessment agreed upon through the consensus process.

As Figure 1.1 illustrates, the themes and periods of U.S. history function as a matrix; the assessment addressed the role of each theme in each of the periods. The framework makes clear that not all themes are equally important in each period. In addition, the

Figure 1.1 NAEP 1994 U.S. History Content Matrix

THEMES / PERIODS	Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	The Gathering and Interactions of Peoples, Cultures, and Ideas	Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	The Changing Role of America in the World
Three Worlds and Their Meeting in the Americas (Beginnings to 1607)				
Colonization, Settlement, and Communities (1607 to 1763)				
The Revolution and the New Nation (1763 to 1815)				
Expansion and Reform (1801 to 1861)				
Crisis of the Union: Civil War and Reconstruction (1850 to 1877)				
The Development of Modern America (1865 to 1920)				
Modern America and the World Wars (1914 to 1945)				
Contemporary America (1945 to Present)				

framework specified the percentage of the assessment that was to address each of the themes and periods at each of the three grade levels, as shown in Tables 1.1 and 1.2. These percentages reflect expectations of the curriculum coverage at each grade level. Because many students are not offered U.S. history in grade 4, the framework also included recommendations for adapting the assessment for fourth-grade students.

The framework was structured to ensure coverage of the historical themes, events, and developments considered most important by the groups that participated in the consensus process, while allowing sufficient flexibility to include creative and thought-provoking exercises that might not fit neatly into a more narrowly defined framework.

Themes in U.S. History. Four historical themes are the core organizing structure of the framework. The themes were intended to ensure that all major branches of historical study were covered and that emphasis on various areas was balanced. The themes are also used to define the subscales for reporting. Table 1.1 contains the percentages of assessment time accorded to each theme by the framework. The themes are as follows:

1. Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

This theme concerns the development of American political democracy from colonial times to the present. It covers political events that shaped American democracy, such as the American Revolution, the Civil War, and the fight for civil rights, as well as the core ideas and principles that underlie our institutions. This theme covers students' knowledge of the founding of the nation, the writing of the Constitution, and other fundamental components of the nation's political history. At the same time, it calls for evaluating students' understanding of the role that major political ideas and conflicts have played at different points in our history.

2. The Gathering and Interactions of Peoples, Cultures, and Ideas

The second theme is broadly defined because it covers a vast component of U.S. history: the interactions among the people and cultures of many countries, racial and ethnic groups, and religious traditions that have contributed to the development of American society. This theme covers the nature and role of immigration throughout our history, cultural developments, patterns of social organization, and changing roles of men and women.

3. Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

This theme focuses on the economic history of the nation and its development from a rural, agricultural society to an urban, industrialized superpower. It covers the roles economic ideas and beliefs have played in this change as well as the roles of geography and of developments in science and technology.

4. The Changing Role of America in the World

This theme calls for coverage of the many factors — physical geography, political ideals, economic interests, public opinion — that have shaped American foreign policy. It also addresses specific interactions between the United States and other nations and the domestic consequences of developments in foreign policy.

Table 1.1 Targeted Distribution of Assessment Time Across Historical Themes

THEMES / GRADE LEVELS	Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	The Gathering and Interactions of Peoples, Cultures, and Ideas	Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	The Changing Role of America in the World
Grade 4	25%	35%	25%	15%
Grade 8	30%	30%	20%	20%
Grade 12	25%	25%	25%	25%

Periods of U.S. History. Eight periods provide chronological structure for the many issues included in the four themes. These periods focus attention on several major eras of U.S. history. They overlap at some points because they were conceived to ensure thorough coverage of major trends and events. The historical periods are not used as reporting subscales, but rather were used in the assessment construction process to ensure appropriate chronological coverage. The periods are as follows:

1. **Three Worlds and Their Meeting in the Americas (Beginnings to 1607)**
2. **Colonization, Settlement, and Communities (1607 to 1763)**
3. **The Revolution and the New Nation (1763 to 1815)**
4. **Expansion and Reform (1801 to 1861)**
5. **Crisis of the Union: Civil War and Reconstruction (1850 to 1877)**
6. **The Development of Modern America (1865 to 1920)**
7. **Modern America and the World Wars (1914 to 1945)**
8. **Contemporary America (1945 to Present)**

The percentages of assessment time devoted to each period, as described in the framework, are presented in Table 1.2.

Ways of Knowing and Thinking About U.S. History. Finally, the framework considers the ways of thinking and kinds of knowledge that historical study requires. These are divided into two general cognitive domains that were used as a guide in exercise development. The two domains and their definitions are as follows:

1. **Historical Knowledge and Perspective**
   This domain includes knowing and understanding people, events, concepts, themes, movements, contexts, and historical sources; sequencing events; recognizing multiple perspectives and seeing an era or movement through the eyes of different groups; and developing a general conceptualization of U.S. history.
2. **Historical Analysis and Interpretation**
   This domain includes explaining issues, identifying historical patterns; establishing cause-and-effect relationships; finding value statements; establishing significance; applying historical knowledge; weighing evidence to draw sound conclusions; making defensible generalizations; and rendering insightful accounts of the past.

Table 1.2 Targeted Distribution of Assessment Time Across Historical Periods

PERIODS / GRADE LEVELS	Beginnings to 1607	1607 to 1763	1763 to 1815	1801 to 1861	1850 to 1877	1865 to 1920	1914 to 1945	1945 to Present
Grade 4	20%	15%	15%	15%	10%	5%	5%	15%
Grade 8	5%	10%	20%	15%	20%	10%	10%	10%
Grade 12	5%	10%	15%	10%	10%	15%	15%	20%

The NAEP 1994 U.S. History Assessment Instruments

In keeping with current assessment research, the framework called for the inclusion of both multiple-choice and constructed-response exercises in the 1994 U.S. history assessment.[5] The framework also required that students be engaged in some of the activities and ways of thinking that are part of the study of history. Thus the assessment included a range of constructed-response (that is, non-multiple-choice) exercises to stimulate thinking and a wide array of stimuli, including texts, maps, photographs, paintings and drawings, political cartoons, advertisements, posters, and graphs, charts, and tables.[6]

Assessment exercises were innovative in other respects. One limitation of many traditional assessments is that they frequently present pieces of information or problems to be solved in isolation. Since students ordinarily consider individual points within the context of the classroom or of material they are reading, assessment questions that appear out of context may not elicit the full range of students' knowledge and abilities.

Two strategies were adopted in the NAEP 1994 U.S. history assessment to address this issue. First, many of the questions appear in sets of two or three, all pertaining to a particular stimulus or issue. In this way, students are encouraged to consider several different aspects of the subject and have a greater opportunity to demonstrate their knowledge and understanding than they otherwise might.

The second strategy was to develop two longer sets of exercises, called *theme blocks*, each focused on a single historical issue. Both theme blocks contain a variety of sources and types of tasks. Students to whom the theme blocks were administered spent all of their assessment time on them. These students had the opportunity to demonstrate their ability to work through many aspects of a historical issue, to work with a variety of primary and secondary sources, to use several different sources at once, and to synthesize a body of information. One theme block was administered at grade 8 and the other at grade 12.

In addition to multiple-choice questions, the assessment included both short and extended constructed-response questions. Each constructed-response question was scored according to a scoring guide, or rubric, that gave credit for partially correct answers. The exercises called for a range of responses. Short constructed-response questions called for a word, a phrase, or a sentence or two to demonstrate understanding of specific material. Extended constructed-response questions called for more developed argument, gathering of evidence, or interpreting of data. Short constructed-response questions were scored according to three-part scoring guides, in which a score of 3 represented an appropriate answer, a score of 2 a partially correct answer, and a score of 1 an inappropriate answer. Extended constructed-response questions were scored according to four-part scoring guides, in which a 4 was assigned to complete responses, a 3 to responses that responded to essential components of the task, a 2 to partially correct responses, and a 1 to inappropriate answers. Many of the constructed-response questions were scaffolded: that is, students were asked to respond first to one portion of the question and then to another. In a few cases, they were asked to do other tasks, such as sequencing events on a time line.

Across the three grades assessed — fourth, eighth, and twelfth — a total of 268 multiple-choice, 96 short constructed-response, and 37 extended constructed-response questions were administered in the 1994 assessment.[7] Some of the questions were administered at more than one grade. Students spent more than half the assessment time responding to constructed-response questions. By grade, the percentage of assessment time spent on constructed-response questions was 57 percent at grade 4, 56 percent at grade 8, and 61 percent at grade 12. Additional information about the assessment design is included in Appendix A.

Description of School and Student Samples

As with all NAEP assessments, the schools and students participating in the 1994 U.S. history assessment were selected through scientifically designed sampling procedures. Approximately 22,000 fourth, eighth, and twelfth graders in 1,500 public and nonpublic schools across the country participated in the assessment.

The results presented in this report are based on representative samples of students at each of the three grade levels. Each school that participated in the assessment and each student assessed represents a portion of the population of interest. As a result, the findings in this report pertain to all fourth, eighth, and twelfth graders in the nation. For a more detailed description of the sample and of the sampling procedures, see Appendix A.

Reporting NAEP U.S. History Results

The NAEP U.S. history assessment provides a great deal of information about the knowledge and abilities of the nation's fourth-, eighth-, and twelfth-grade students. To maximize the usefulness of this information to policymakers, educators, parents, and other interested parties, the NAEP results are presented not only as average scores on the U.S. history scale, but also in terms of the percentages of students attaining NAEP U.S. history achievement levels. Thus, NAEP results provide information about what students *know and can do*, as well as indicate the extent to which their achievement meets expectations of what students *should know and be able to do.*

U.S. History Scale

Student responses to the NAEP 1994 U.S. history assessment were analyzed to determine the percentages of students responding correctly to each multiple-choice question and the percentage of students performing in each of the score categories for constructed-response questions. Item response theory (IRT) methods were used to produce within-grade scales that summarize results for each of the four themes described earlier. Each subscale for grade 4 was linked to the corresponding subscale for grade 8. Likewise, each subscale for grade 12 was linked to the corresponding subscale for grade 8. Then, each linked subscale was mapped onto a 0 to 500 scale. These separate subscales were then weighted by the percentages shown in Table 1.1 to produce a composite NAEP U.S. history scale. Chapters 2, 4, and 5 present results based on the U.S. history scale. (Please note that the scales for each of the NAEP subjects assessed in 1994 — reading, U.S. history, and geography — were developed independently. Therefore, results should not be compared across subjects. Details of the scaling procedures are presented in the *NAEP 1994 Technical Report.*)

U.S. History Achievement Levels

The 1994 assessment results are also reported using the U.S. history achievement levels that were authorized by the NAEP legislation and adopted by the National Assessment Governing Board. The achievement levels are based on collective judgments about what students *should know and be able to do* relative to the body of content reflected in the NAEP assessment framework. Three achievement levels were defined for each grade level assessed: *Basic, Proficient,* and *Advanced.* The levels were defined by a broadly representative panel of teachers, education specialists, and members of the general public.

For reporting purposes, the achievement levels for each grade are placed on the NAEP U.S. history scale defining four ranges — *Basic, Proficient, Advanced,* and below *Basic.* The policy definitions of the three achievement levels are presented in Figure 1.2.

Figure 1.2 Policy Definitions of NAEP Achievement Levels

Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.
Proficient	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.
Advanced	This level signifies superior performance.

It should be noted that the setting of achievement levels for the national assessment is relatively new and in transition. Some evaluations have concluded that the percentages of students at certain levels may be underestimated.[8] On the other hand, critiques of those evaluations have found that such conclusions are not supported by the weight of the empirical evidence.[9]

The student achievement levels in this report have been developed carefully and responsibly and have been subject to refinements and revisions in procedures as new technologies have become available. Upon reviewing the available information, the Commissioner of NCES has judged that the achievement levels are in a developmental status. However, the Commissioner and the Governing Board also believe that the achievement levels are useful and valuable in reporting on the educational achievement of students in the United States. Results reported in terms of the U.S. history achievement levels are presented in Chapter 3 of this report.

Example Questions from the NAEP 1994 U.S. History Assessment

As discussed earlier in this chapter, the NAEP 1994 U.S. history assessment is a rich collection of exercises developed to survey the historical knowledge and skills of students in grades 4, 8, and 12. Each student received a mixture of multiple-choice and constructed-response questions.

Figure 1.3 presents examples of assessment exercises. (Additional example questions are included in Appendix C.) The tables accompanying the exercises present two types of percentages: (1) the overall percentage of students within a grade who successfully answered the question, and (2) the percentage of students within each of the achievement level intervals — *Basic*, *Proficient*, *Advanced*, and below *Basic* — who successfully answered the questions. (A fuller description of the achievement levels at each grade can be found in Chapter 3. Note also that at all three grades, the percentages for students within the *Advanced* achievement level interval are not presented because of small sample sizes.)

The first exercise in Figure 1.3 is a multiple-choice question administered at grade 4, based on an excerpt from a Martin Luther King Jr., speech. The table shows that the majority (87 percent) of all fourth graders correctly answered this question. Seventy-two percent of fourth graders who scored below the *Basic* achievement level answered the question correctly, compared to 93 percent for students who scored within the *Basic* level. Almost all the fourth graders (99 percent) who scored within the *Proficient* achievement level interval answered the question correctly.

The second exercise shown in Figure 1.3 is a short constructed-response question administered at grade 8. The stimulus material for this exercise is a reproduction of an historical print depicting the Boston Tea Party. The question is "scaffolded," meaning that students are first asked to identify the event depicted and then to describe the importance of the event. A sample response for a student who received a score of "Appropriate" (as defined by the scoring rubic) on this question is also provided. As the table accompanying this exercise shows, almost one-fifth (19 percent) of all eighth graders wrote answers rated as "Appropriate." For those eighth graders who scored below the *Basic* achievement level, only 2 percent wrote answers rated as "Appropriate," compared to 20 percent within the *Basic* level. Approximately two-thirds (64 percent) of the eighth graders who scored within the *Proficient* achievement level interval wrote responses rated as "Appropriate."

The third exercise shown in Figure 1.3 is an extended constructed-response question administered at grade 12. The question was part of a grade 12 theme block on the Depression and required students to integrate information from a variety of sources. (A description of the NAEP U.S. history theme blocks appears earlier in this chapter.) Sample responses for students who received scores of "Essential" and "Complete" are provided. For this extended constructed-response question, one-fifth (20 percent) of all twelfth graders provided answers rated "Essential" or better. For those twelfth graders who scored below the *Basic* achievement level, 6 percent provided answers rated "Essential" or better, compared with 27 percent of students within the *Basic* level. Almost two-thirds (64 percent) of twelfth graders who scored within the *Proficient* achievement level interval provided answers rated "Essential" or better.

Figure 1.3 NAEP 1994 U.S. History Sample Questions

Example of a Grade 4 Multiple-Choice Question

I have a dream that one day this nation will rise up and live out the true meaning of its creed: "We hold these truths to be self-evident; that all men are created equal."

I have a dream that one day on the red hills of Georgia the sons of former slaves and the sons of former slaveowners will be able to sit down together at the table of brotherhood.

I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character.

I have a dream that one day... little Black boys and Black girls will be able to join hands with little White boys and girls and walk together as sisters and brothers.

The speech was given by

A Abraham Lincoln

B Gloria Steinem

C George Bush

▶ D Martin Luther King Jr.

Historical Theme:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

Historical Period:
Contemporary America (1945 to Present)

Grade 4	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	*Below Basic* 194 and below*	*Basic* 195-242*	*Proficient* 243-275*	*Advanced* 276 and above*
87 (1.0)	72 (2.6)	93 (1.2)	99 (0.7)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (see Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Figure 1.3 NAEP 1994 U.S. History Sample Questions (continued)

Example of a Grade 8 Short Constructed-Response Question

THE DESTRUCTION OF TEA AT BOSTON HARBOR

The Destruction of Tea at Boston Harbor, N. Currier 1846. Museum of the City of New York. The Harry T. Peters Collection.

Identify the event that is portrayed in the picture above.

Why is the event important in United States history?

Historical Theme:

Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

Historical Period:

The Revolution and the New Nation (1763 to 1815)

Responses to this question were scored according to a three-level rubric as 1) Inappropriate, 2) Partial, and 3) Appropriate.

Grade 8	Percentage "Appropriate" within Achievement Level Intervals			
Overall Percentage Appropriate	***Below Basic*** **251 and below***	***Basic*** **252-293***	***Proficient*** **294-326***	***Advanced*** **327 and above***
19 (1.2)	**2 (0.7)**	**20 (2.1)**	**64 (4.1)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (s ppendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Figure 1.3 NAEP 1994 U.S. History Sample Questions (continued)

Sample Response (Score of 3)

Identify the event that is portrayed in the picture above.

The Boston Tea Party

Why is the event important in United States history?

It was a rebellion against British Rule in particular against the British monopoly on tea that preceded the American Revolution

An **Appropriate** response (score of 3) correctly identifies the event as the Boston Tea Party and explains that the event is important because it showed colonial resistance to British policies, that in turn led to the American Revolution.

Figure 1.3 NAEP 1994 U.S. History Sample Questions (continued)

Example of a Grade 12 Extended Constructed-Response Question

The next three documents (Documents J-L) were produced by people who believed either that the New Deal had done too much or that it was not doing enough to end the Depression.

Document J - *Portion of a campaign song for Senator Huey Long of Louisiana, 1935*

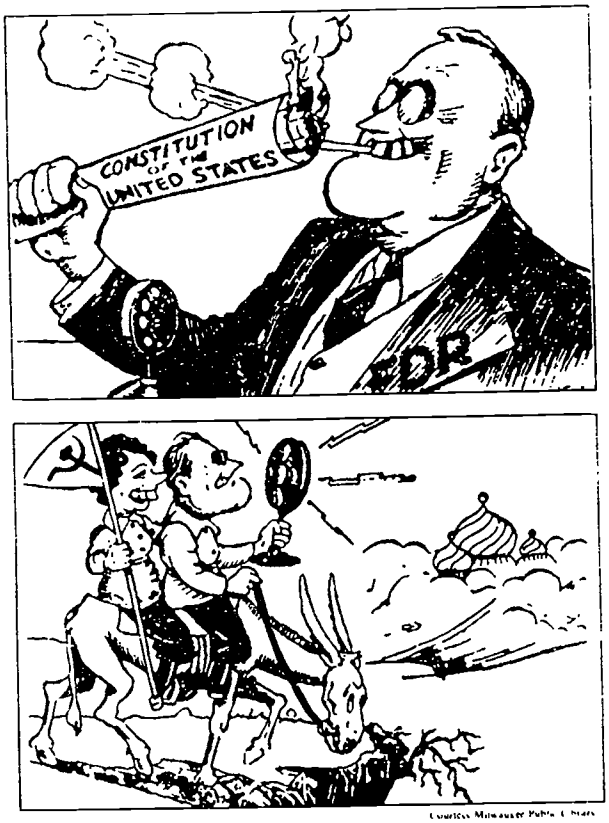
Every Man a King

Why weep or slumber America
Land of brave and true
With castles and clothing and food for all
All belongs to you
Ev'ry man a King, ev'ry man a King
For you can be a millionaire
But there's something belonging to others
There's enough for all people to share

Document K - *Excerpt from a speech given by Huey Long, March 7, 1935*

So it has been while millions have starved and gone naked and while babies have cried and died wanting milk; so it has been while people begged for meat and bread to eat. Mr. Roosevelt's administration has sailed merrily along, plowing under and destroying the things to eat and wear, with tear-dimmed eyes and hungry souls made to chant for this New Deal so that even their starvation dole is not taken away from them, and meanwhile the food and clothes craved by humanity for their bodies and souls go to destruction and ruin.

Document L - *Two Anti-Roosevelt Cartoons*



What criticisms of Roosevelt are being made in the cartoons?

Identify the ways that the cartoons differ from the Huey Long passages (Documents J-K) in their criticisms of Roosevelt.

Historical Theme:

Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

Historical Period:

Modern America and the World Wars (1914 to 1945)

Responses to this question were scored according to a four-level rubric as 1) Inappropriate, 2) Partial, 3) Essential, and 4) Complete.

Grade 12	Percentage "Essential" or Better within Achievement Level Intervals			
Overall Percentage Essential or Better	***Below Basic*** **293 and below***	***Basic*** **294-324***	***Proficient*** **325-354***	***Advanced*** **355 and above***
20 (1.2)	**6 (1.2)**	**27 (3.1)**	**64 (6.5)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (see Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Figure 1.3 NAEP 1994 U.S. History Sample Questions (continued)

Sample Response (Score of 3)

What criticisms of Roosevelt are being made in the cartoons?

The criticisms that are being made in the cartoons are that Roosevelt is ignoring people's rights and what the constitution is all about.

Identify the ways that the cartoons differ from the Huey Long passages (Documents J-K) in their criticisms of Roosevelt.

The cartoons differ from the Huey Long passages in that the cartoons focus on the constitution and Foreign affairs, and Long focused on specifically the economic problems at home including unemployment.

An **Essential** response (score of 3) explains the criticisms of Roosevelt made in the cartoons, and fails to offer or offers a general comparison of the cartoons and Long documents.

Figure 1.3 NAEP 1994 U.S. History Sample Questions (continued)

Sample Response (Score of 4)

What criticisms of Roosevelt are being made in the cartoons?

The ~~a~~ cartoons beeing made are that roosevelt is overstepping the limits of his powers as president and violating the constution And that he has communistic and socialistic trends in his programs and policies,

Identify the ways that the cartoons differ from the Huey Long passages (Documents J-K) in their criticisms of Roosevelt.

Huey long believed that Roosevelt was falling short where as these cartoons indicate he is going to far

A **Complete** response (score of 4) correctly explains the criticisms of Roosevelt made in the cartoons, and compares the cartoons with the left-wing criticisms made in the Long documents.

Endnotes

1. Anonymous teacher quoted in "Engaging the Past," Alice Almond Shrock and Randall Shrock. (1994). *Journal of American History,* 81(3), 1093-98.
2. *The National Educational Goals Report: Building a Nation of Learners.* (1994). (Report of the National Education Goals Panels: U.S. Government Printing Office).
3. *Social studies education: Kindergarten-grade 12. National survey.* Council of State Social Studies Specialists (1986). (ERIC Document Reproduction Service No. 289800).
4. *U.S. History framework for the 1994 National Assessment of Educational Progress.* (Washington, DC: National Assessment Governing Board).
5. Cornbleth, Catherine. (1986). *Assessing skills and thinking in social studies.* Study Group on the National Assessment of Student Achievement. (ERIC Document Reproduction Service No. ED 279671).

   Fraenkel, Jack R. (1986). *The assessment of social studies knowledge* [Draft]. Study Group on the National assessment of Student Achievement. (ERIC Document Reproduction Service No. ED 279676).

   Nickell, Pat. (1993). *Alternative assessment: Implications for social studies.* ERIC Clearinghouse for Social Studies/Social Science Education. (ERIC Document Reproduction Service No. ED 360 219).

   Wiggins, Grant. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan,* 70(9), 703-713.
6. *NCSS Position Statement on Testing and Evaluation of Social Studies Students: Position Statement and Guidelines.* (National Council for the Social Studies, Washington, DC, 1991).
7. NAEP assessments are organized in a "BIB-spiral" design, which means that each individual student receives only a subset of the total pool of questions for his or her grade. The questions are arranged in "blocks," which are assigned to booklets in a way that ensures that each student will complete 50 minutes worth of questions, and that the blocks will be distributed throughout the assessment population. Thus, complete coverage of the content is possible without overburdening students and schools.
8. *Education achievement standards: NAGB's approach yields misleading interpretations.* (1993). (United States General Accounting Office Report to Congressional Requestors. Washington, DC: United States General Accounting Office).

   Setting Performance Standards for Student Achievement. A Report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment: An Evaluation of the 1992 Achievement Levels. (Stanford, CA: National Academy of Education, 1993).
9. *Technical report on setting achievement levels on the 1992 National Assessment of Educational Progress in mathematics, reading, and writing.* (1993). American College Testing (Washington, DC: National Assessment Governing Board).

   Cizek, G., (1993). *Reactions to National Academy of Education report.* (Washington, DC: National Assessment Governing Board).

   Kane, M., (1993). *Comments on the NAEP evaluation of the NAGB achievement levels.* (Washington, DC: National Assessment Governing Board).

   NAEP Reading revisited: An evaluation of the 1992 achievement levels descriptions. (1995). American College Testing (Washington, DC: National Assessment Governing Board).

CHAPTER 2

U.S. History Results for the Nation and Regions

This chapter presents the NAEP U.S. history scale scores of students in grades 4, 8, and 12, reported using a scale that ranges from 0 to 500. Findings are presented for the nation, for regions of the country, and for selected subgroups of students. (For a visual representation of student performance on each thematic subscale — also ranging from 0 to 500 — see pages 66-71 in Chapter 5 of this report.)

The results provided in this chapter address statistically significant differences that were found between reporting subgroups. In other cases, score estimates for various subgroups may appear to differ, but these differences are not statistically significant. (Significant differences are those that are unlikely to be due to sampling variability or chance.)

In addition, the NAEP 1994 U.S. history assessment results are explored in more depth by examining the interactions among several major reporting variables. Average U.S. history scale scores are examined for subgroups of students within various demographic populations. By doing so, it is possible to determine if general patterns of U.S. history performance for certain groups of students are related to additional background characteristics.

U.S. History Results for the Nation

Table 2.1 presents the average U.S. history scale scores of students nationwide.

- The average score of students in grade 4 was 205. The bottom 10 percent of the population scored at or below 147 and the top 10 percent scored at or above 253.
- At grade 8, the average score was 259. The bottom 10 percent of the population scored at or below 217, while the top 10 percent scored at or above 299.
- The average score of students in grade 12 was 286. The bottom 10 percent scored at or below 243 and the top 10 percent scored at or above 326.

TABLE 2.1

Average U.S. History Scale Scores by Percentile Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

Grade Levels	Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Grade 4	205 (1.0)	147 (2.4)	180 (1.3)	210 (0.9)	234 (1.2)	253 (1.2)
Grade 8	259 (0.6)	217 (1.0)	239 (0.9)	261 (0.8)	282 (0.8)	299 (0.8)
Grade 12	286 (0.8)	243 (1.0)	265 (0.9)	288 (0.8)	309 (1.0)	326 (0.8)

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Figure 2.1 and Table 2.2 present results for the nation as well as the four regions of the country: Northeast, Southeast, Central, and West. The composition of the regions is described in Appendix A. At all three grades, students in the Central region had higher average scores than did students in the Southeast. At grade 4, students from the Central region also outperformed students from the West.

At grade 8, several other differences were significant. Students from the Southeast were outperformed by students from the Northeast and the West. Also, eighth graders from the West were outperformed by their counterparts from the Central and Northeast regions.

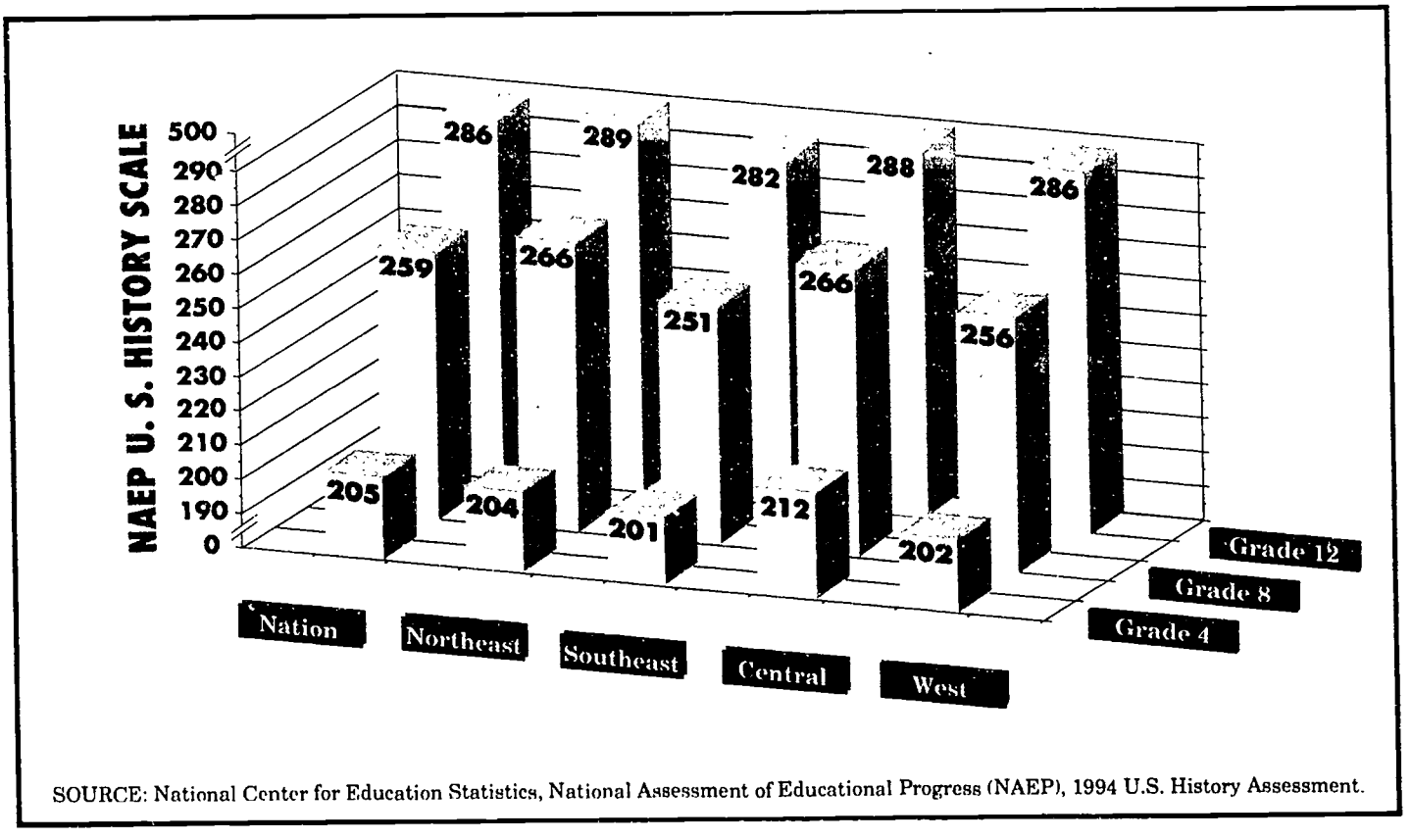
At grade 12, students from the Northeast had higher average scores than did those from the Southeast.

U.S. History Results for Major Reporting Subgroups

Tables 2.3 through 2.8 present the average U.S. history scale scores for major subgroups of the fourth-, eighth-, and twelfth-grade student populations.

Cautions in Interpretations. In viewing the results presented in this section, the reader is cautioned against making simple or causal inferences about subgroup membership or about the effectiveness of Title I programs or public and nonpublic schools. Average performance differences between groups of students may result from socioeconomic, home background, or other factors. For example, performance differences observed among racial/ethnic subgroups are almost certainly associated with a broad range of socioeconomic and educational variables. Similarly, performance differences between public and nonpublic school students may be better understood after factors such as composition of the student body, parents' education levels, and parental involvement are considered.

Figure 2.1 Average NAEP U.S. History Scale Scores by Grade and by Region



SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

TABLE 2.2

Average U.S. History Scale Scores for the Nation and by Region Grades 4, 8, and 12

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Region		
Northeast	22 (0.7)	204 (2.4)
Southeast	23 (1.0)	201 (1.9)
Central	25 (0.8)	212 (2.6)
West	30 (0.6)	202 (2.1)
Grade 8		
Nation	100	259 (0.6)
Region		
Northeast	20 (0.8)	266 (1.7)
Southeast	25 (0.9)	251 (1.3)
Central	24 (0.6)	266 (1.3)
West	31 (0.8)	256 (1.1)
Grade 12		
Nation	100	286 (0.8)
Region		
Northeast	20 (0.5)	289 (1.9)
Southeast	23 (0.8)	282 (1.4)
Central	27 (0.7)	288 (1.4)
West	30 (0.7)	286 (1.6)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Race/Ethnicity. Results are presented for students in different racial/ethnic groups based on the students' self-identification of their race/ethnicity according to the following mutually exclusive categories: White, Black, Hispanic, Asian, Pacific Islander, and American Indian (including Alaskan Native). Average scale scores for students in various racial and ethnic groups are shown in Table 2.3. The performance of these groups varied significantly. Evident at all three grades was the finding that White and Asian students had higher average U.S. history scores than did Black and Hispanic students.

At the fourth grade, Pacific Islander students had higher scores than did Hispanic or Black students. Also, White students had a higher average history score than did American Indian students.

TABLE 2.3

Average U.S. History Scale Scores by Race/Ethnicity Grades 4, 8, and 12

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Race/Ethnicity		
White	69 (0.3)	215 (1.2)
Black	15 (0.1)	177 (1.6)
Hispanic	11 (0.2)	180 (2.7)
Asian	2 (0.2)	209 (4.6)
Pacific Islander	1 (0.2)	200 (5.9)
American Indian	2 (0.3)	190 (6.1)
Grade 8		
Nation	100	259 (0.6)
Race/Ethnicity		
White	69 (0.2)	267 (0.8)
Black	15 (0.1)	239 (1.4)
Hispanic	11 (0.1)	243 (1.3)
Asian	2 (0.1)	270 (3.6)
Pacific Islander	1 (0.3) !	252 (7.1) !
American Indian	1 (0.3) !	246 (3.7) !
Grade 12		
Nation	100	286 (0.8)
Race/Ethnicity		
White	74 (0.4)	292 (0.8)
Black	12 (0.2)	265 (1.5)
Hispanic	9 (0.3)	267 (1.6)
Asian	3 (0.2)	287 (4.0)
Pacific Islander	1 (0.2)	280 (3.9)
American Indian	1 (0.2) !	279 (4.0) !

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding or, in the case of the race/ethnicity variable, because some students categorized themselves as "other."

! Interpret with caution any comparisons involving this statistic. The nature of the sample does not allow for accurate determination of the variability of this value.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

At grade 12, Pacific Islander students had higher scores than did Hispanic or Black students. In addition, White students outperformed Pacific Islander students.

For the American Indian student samples at grades 8 and 12, and for the Pacific Islander student sample at grade 8, the nature of the samples does not allow accurate determination of the standard errors. For this reason, differences among these samples and other racial/ethnic subgroups are not discussed.

Gender. Scale score results by gender are presented in Table 2.4. Only one statistically significant difference was found in these data: at grade 12, males had higher average U.S. history scores than did females.

TABLE 2.4

Average U.S. History Scale Scores by Gender Grades 4, 8, and 12

THE NATION'S REPORT CARD NAEP 1994 U.S. History Assessment

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Gender		
Male	50 (0.8)	203 (1.5)
Female	50 (0.8)	206 (1.1)
Grade 8		
Nation	100	259 (0.6)
Gender		
Male	50 (0.5)	259 (0.8)
Female	50 (0.5)	259 (0.7)
Grade 12		
Nation	100	286 (0.8)
Gender		
Male	50 (0.8)	288 (0.8)
Female	50 (0.8)	285 (0.9)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Parents' Level of Education. Students were asked to identify the level of education completed by their parents or guardians: did not finish high school, graduated from high school, had some education after high school, or graduated from college. (Those who did not have this information chose the response option "I don't know.") The highest level reported for either parent was used in this analysis.

The level of education that a parent had reportedly completed shows a consistent relationship to student performance on the assessment. Table 2.5 presents these results. It is important to note in reviewing these results, however, that 34 percent of fourth graders and 9 percent of eighth graders reported that they did not know what level of education their parents had completed. Furthermore, some existing research has questioned the accuracy of student-reported data among similar groups of students.[1] Nevertheless, across NAEP assessments, increasing levels of parents' education corresponded with higher average scores.

At all three grades, each additional increment of reported parental education corresponds to a significantly higher level of student performance. The only exception to this pattern occurred at grade 4. There was no significant difference in performance between fourth graders who reported that a parent had completed college and those who reported that a parent completed some education after high school.

TABLE 2.5

Average U.S. History Scale Scores by Parents' Highest Education Level Grades 4, 8, and 12

THE NATION'S REPORT CARD NAEP 1994 U.S. History Assessment

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Parents' Education Level		
Graduated College	43 (0.9)	216 (1.2)
Some Education After High School	8 (0.4)	214 (2.1)
Graduated High School	12 (0.5)	197 (1.8)
Did Not Finish High School	4 (0.3)	177 (3.3)
I Don't Know	34 (0.8)	195 (1.4)
Grade 8		
Nation	100	259 (0.6)
Parents' Education Level		
Graduated College	42 (1.0)	270 (0.8)
Some Education After High School	19 (0.5)	264 (0.8)
Graduated High School	23 (0.8)	251 (0.8)
Did Not Finish High School	7 (0.4)	241 (1.3)
I Don't Know	9 (0.4)	238 (1.4)
Grade 12		
Nation	100	286 (0.8)
Parents' Education Level		
Graduated College	45 (1.0)	296 (0.9)
Some Education After High School	25 (0.7)	287 (1.2)
Graduated High School	20 (0.7)	276 (1.1)
Did Not Finish High School	7 (0.4)	263 (1.4)
I Don't Know	3 (0.2)	256 (2.7)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Type of Location. Table 2.6 presents the average scores of students attending schools in each of three major types of locations: central city, urban fringe/large town, and rural/small town. These categories indicate the geographic locations of the schools and are not intended to indicate or imply social or economic meanings for location types. These categories are based on U.S. Bureau of the Census definitions of standard metropolitan statistical areas, population size, and density. (The type of location classifications are described in Appendix A.)

The results for grades 4 and 8 show the same pattern, with students attending school in urban fringe/large towns having higher average scores than students attending either central city or rural/small-town schools. At grade 12 the pattern was slightly different. On average, both students attending urban fringe/large town and those attending central city schools outperformed students attending rural/small-town schools.

Title I Participation. Staff members at each school that took part in the NAEP 1994 U.S. history assessment were asked to identify which of the students participated in Title I programs or received services funded by Title I grants.[2] The Title I legislation provides funds to state and local educational agencies to support programs aimed at assisting disadvantaged students (those who are failing or are at risk of failing) in low-income communities.

Table 2.7 presents the U.S. history assessment results for students who received Title I services and for those who did not. Differences in performance between these recipients and nonrecipients should not be viewed as indicative of the success or failure of Title I programs. Title I services are intended for students who typically score poorly on assessments.

As can be seen from the NAEP 1994 U.S. history results, the percentage of students receiving Title I

TABLE 2.6

Average U.S. History Scale Scores by Type of Location Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Type of Location		
Central City	35 (2.1)	198 (2.0)
Urban Fringe/Large Town	43 (2.4)	211 (1.6)
Rural/Small Town	22 (2.1)	203 (2.8)
Grade 8		
Nation	100	259 (0.6)
Type of Location		
Central City	36 (2.4)	257 (1.3)
Urban Fringe/Large Town	38 (2.9)	262 (1.2)
Rural/Small Town	26 (1.8)	258 (1.3)
Grade 12		
Nation	100	286 (0.8)
Type of Location		
Central City	31 (2.2)	286 (1.3)
Urban Fringe/Large Town	43 (2.7)	289 (1.2)
Rural/Small Town	26 (2.0)	281 (1.0)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

TABLE 2.7

Average U.S. History Scale Scores by Title I Participation Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Title I Participation		
Yes	14 (1.3)	162 (2.2)
No	86 (1.3)	211 (1.1)
Grade 8		
Nation	100	259 (0.6)
Title I Participation		
Yes	7 (1.0)	232 (1.8)
No	93 (1.0)	261 (0.7)
Grade 12		
Nation	100	286 (0.8)
Title I Participation		
Yes	2 (0.5) !	254 (2.1) !
No	98 (0.5)	287 (0.8)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

! Interpret with caution any comparisons involving this statistic. The nature of the sample does not allow for accurate determination of the variability of this value.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

services is greatest in the elementary grades (14 percent at grade 4) and decreases as students progress through middle school (7 percent at grade 8). At grades 4 and 8, the average scale score of students receiving Title I services was significantly lower than that of students not receiving Title I services. (Differences for twelfth graders are not discussed here because the nature of the grade 12 sample does not allow for accurate estimation of the variability of the percentages of Title I recipients.)

Type of School. The average scale scores by type of school are presented in Table 2.8. At all three grades, students attending nonpublic schools outperformed public school students.

TABLE 2.8 **Average U.S. History Scale Scores by Type of School Grades 4, 8, and 12**

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage of Students	Average Scale Score
Grade 4		
Nation	100	205 (1.0)
Type of School		
Public Schools	90 (0.8)	203 (1.2)
Nonpublic Schools	10 (0.8)	222 (1.9)
Catholic Schools	6 (0.7)	221 (2.5)
Other Nonpublic Schools	4 (0.5)	224 (3.1)
Grade 8		
Nation	100	259 (0.6)
Type of School		
Public Schools	90 (0.9)	257 (0.7)
Nonpublic Schools	10 (0.9)	278 (1.1)
Catholic Schools	6 (0.6)	279 (1.5)
Other Nonpublic Schools	4 (0.6)	277 (2.1)
Grade 12		
Nation	100	286 (0.8)
Type of School		
Public Schools	89 (1.1)	284 (0.8)
Nonpublic Schools	11 (1.1)	299 (1.3)
Catholic Schools	6 (0.9)	298 (2.2)
Other Nonpublic Schools	5 (0.6)	299 (2.2)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

An In-Depth Look at Selected Background Characteristics

One way to take a closer look at the performance of students within selected demographic populations is to see if the magnitude of the differences between groups of students varies when other background characteristics are taken into account. This section presents NAEP 1994 U.S. history results for subgroups of students within various demographic populations. Three specific background characteristics are explored: gender, race/ethnicity, and parents' highest level of education. In addition, type of school (public and nonpublic) results conditioned on parents' education are examined.

Gender and Race/Ethnicity. As reported earlier in this chapter, the average U.S. history scores for males and females at grades 4 and 8 tended to be similar. Only at grade 12 was there a significant difference, with males outperforming females. At all grades, White students displayed higher average U.S. history scores than Black or Hispanic students. (Asian, Pacific Islander, and American Indian students are not included in this section because, for the purposes of this analysis, their sample sizes are insufficient.) One question that might be asked is whether or not the pattern in male and female students' scores holds regardless of race/ethnicity. Was the difference in performance between male and female students larger in some racial/ethnic subgroups than in others?

Table 2.9 presents results of analyses carried out to answer these questions. Average scores of male and female students and the differences between these averages are presented separately for three racial/ethnic subgroups. As displayed in the table, the overall grade 12 difference between male and female students was significant only for Hispanic students. For all three racial/ethnic groups, the average scores for male students at grade 12 were higher than those of females, but only the 9 point difference for Hispanic students was significant. In addition, for Black fourth graders (unlike the overall results at grade 4 that indicated no significant difference between male and female students) female students outperformed male students.

A comparison of the magnitude of the score differences between male and female students yielded little or no evidence that these differences vary significantly across racial/ethnic groups of students. The only significant difference is at grade 4, where the male/female difference for Black students is greater than that for White students.

TABLE 2.9

Average U.S. History Scale Scores of Male and Female Students by Race/Ethnicity Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Overall	White	Black	Hispanic
Grade 4				
Female	206 (1.1)	215 (1.1)	183 (2.4)	182 (3.3)
Male	203 (1.5)	215 (1.8)	170 (1.9)	178 (3.4)
Female - Male =	***2 (1.9)***	***0 (2.1)***	***13 (3.1)****	***3 (4.7)***
Grade 8				
Female	259 (0.7)	266 (0.8)	240 (2.1)	243 (2.2)
Male	259 (0.8)	267 (1.0)	236 (1.5)	243 (1.5)
Female - Male =	***0 (1.0)***	***-1 (1.3)***	***4 (2.6)***	***0 (2.7)***
Grade 12				
Female	285 (0.9)	291 (1.0)	263 (1.7)	262 (1.7)
Male	288 (0.8)	293 (0.9)	267 (2.1)	271 (2.4)
Female - Male =	***-3 (1.2)****	***-2 (1.3)***	***-4 (2.6)***	***-9 (2.9)****

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the average scale scores and the standard errors of the differences appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Differences were calculated prior to rounding.

* Indicates either an overall significant difference, or a significant difference between gender subgroups of students within a specified racial/ethnic subgroup.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Gender and Parental Education. An analysis of gender differences in U.S. history performance in relation to parents' highest level of education is presented in Table 2.10. Average scores of male and female students and the differences between these averages are presented separately for the different levels of parental education reported by students. Were differences in average U.S. history scores between male and female students evident for students at all levels of parental education? Further, were gender differences greater at some parental education levels than at others?

No significant differences in the performance of male and female fourth and eighth graders were found at any level of parental education. Consistent with the overall gender results, average scores for male twelfth graders were higher than those for female twelfth graders across all levels of parental education. However, unlike the overall results, none of these male/female differences were significant.

TABLE 2.10

Average U.S. History Scale Scores of Male and Female Students in Relation to Parents' Highest Education Level Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994

	Overall	Less than High School	Graduated High School	Some Education after High School	Graduated College
Grade 4					
Female	206 (1.1)	181 (5.3)	200 (2.6)	218 (2.6)	216 (1.3)
Male	203 (1.5)	173 (4.5)	194 (3.0)	211 (3.7)	215 (1.8)
Female - Male =	***2 (1.9)***	***9 (7.0)***	***6 (4.0)***	***7 (4.6)***	***1 (2.2)***
Grade 8					
Female	259 (0.7)	242 (1.7)	250 (1.0)	265 (1.1)	270 (0.9)
Male	259 (0.8)	240 (2.2)	251 (1.3)	263 (1.3)	270 (1.0)
Female - Male =	***0 (1.0)***	***2 (2.8)***	***-2 (1.6)***	***2 (1.7)***	***0 (1.3)***
Grade 12					
Female	285 (0.9)	262 (1.8)	274 (1.4)	285 (1.5)	295 (1.2)
Male	288 (0.8)	265 (1.8)	278 (1.6)	288 (1.4)	297 (0.8)
Female - Male =	***-3 (1.2)****	***-3 (2.6)***	***-4 (2.1)***	***-3 (2.1)***	***-2 (1.4)***

Differences between the groups may be partially explained by other factors not included in this table.

Please note that 34 percent of fourth graders and 9 percent of eighth graders reported that they did not know what level of education their parents had completed.

The standard errors of the average scale scores and the standard errors of the differences appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Differences were calculated prior to rounding

* Indicates either an overall significant difference, or a significant difference between gender subgroups of students within a specified level of parents' education.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Type of School (Public and Nonpublic) and Parental Education. As stated earlier in this chapter, students attending nonpublic schools had significantly higher average scores at all grades than did those attending public schools. Is this difference in U.S. history performance between nonpublic and public school students evident at all reported levels of parental education? Further, was the difference in performance greater for some levels of parental education than for others?

Table 2.11 presents results of analyses carried out to address these questions. One striking finding is that the sample sizes for students attending nonpublic schools whose parents did not graduate high school were too small to estimate an average score. Therefore, discussions of public/nonpublic differences at this level of parental education are not possible. As with the overall results, students attending nonpublic schools in each grade had higher average scores than their public school counterparts across all levels of parental education. The exception was fourth graders who reported that at least one parent had some education after high school. Here, the difference (though consistent with the pattern of differences within other groups) was not significant.

A comparison of the magnitude of the score differences between nonpublic and public school students yielded little or no evidence that these differences varied significantly across levels of parental education.

TABLE 2.11

Average U.S. History Scale Scores of Students Attending Public and Nonpublic Schools in Relation to Parents' Highest Education Level Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Overall	Less than High School	Graduated High School	Some Education after High School	Graduated College
Grade 4					
Nonpublic	222 (1.9)	***	212 (4.2)	223 (4.4)	229 (2.3)
Public	203 (1.2)	176 (3.5)	196 (1.9)	213 (2.3)	214 (1.5)
Nonpublic - Public	***19 (2.2)****	***	***16 (4.6)****	***10 (5.0)***	***15 (2.7)****
Grade 8					
Nonpublic	278 (1.1)	***	267 (1.8)	276 (2.3)	282 (1.4)
Public	257 (0.7)	240 (1.3)	249 (0.9)	263 (0.8)	268 (0.9)
Nonpublic - Public	***21 (1.3)****	***	***18 (2.0)****	***13 (2.5)****	***14 (1.7)****
Grade 12					
Nonpublic	299 (1.3)	***	289 (2.1)	293 (1.6)	305 (1.2)
Public	284 (0.8)	262 (1.5)	275 (1.1)	286 (1.3)	295 (1.0)
Nonpublic - Public	***14 (1.5)****	***	***14 (2.4)****	***7 (2.0)****	***10 (1.5)****

Differences between the groups may be partially explained by other factors not included in this table.

Please note that 34 percent of fourth graders and 9 percent of eighth graders reported that they did not know what level of education their parents had completed.

The standard errors of the average scale scores and the standard errors of the differences appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Differences were calculated prior to rounding.

* Indicates either an overall significant difference, or a significant difference between public and nonpublic subgroups of students within a specified level of parents' education.

*** Sample size insufficient to permit a reliable estimate (see Appendix A)

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Race/Ethnicity and Parental Education. Racial/ethnic differences in twelfth-grade average U.S. history scores are presented in relation to parental education levels in Table 2.12. The average scores of White, Black, and Hispanic students and the differences between those scores are presented separately for the different levels of parents' education reported by students. One question that can be answered with these data is whether or not the differences in average U.S. history scores among White, Black, and Hispanic students were evident for students at all levels of parental education. Also, were the performance differences between racial/ethnic groups greater at some parental education levels than at others?

TABLE 2.12

Average U.S. History Scale Scores of White, Black, and Hispanic Students in Relation to Parents' Highest Education Level Grade 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Overall	Less than High School	Graduated High School	Some Education after High School	Graduated College
Grade 12					
White	292 (0.8)	271 (2.3)	281 (1.3)	291 (1.3)	300 (0.9)
Black	265 (1.5)	251 (3.0)	258 (2.3)	269 (2.2)	273 (2.1)
Hispanic	267 (1.6)	256 (2.2)	264 (3.6)	277 (2.4)	277 (2.4)
White - Black =	***27 (1.7)****	***20 (3.8)****	***23 (2.6)****	***22 (2.5)****	***27 (2.3)****
White - Hispanic =	***26 (1.8)****	***14 (3.2)****	***17 (3.9)****	***14 (2.7)****	***23 (2.6)****
Black - Hispanic =	***-2 (2.2)***	***-6 (3.7)***	***-7 (4.3)***	***-8 (3.2)****	***-4 (3.2)***

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the average scale scores and the standard errors of the differences appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Differences were calculated prior to rounding.

* Indicates either an overall significant difference, or a significant difference between racial/ethnic subgroups of students within a specified level of parents' education.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

At the twelfth grade, the average scores of White students were significantly higher than that of Black or Hispanic students across all levels of parental education. In addition, Hispanic twelfth graders whose parents had some education after high school outperformed Black students whose parents had some education after high school. (Data for fourth and eighth graders are not presented in this tabulation because of wide variation among these groups in the accuracy of reporting parental education. See discussion in Appendix A under "Parents' Education Level" for further details.)

At the twelfth grade, the magnitude of the difference in U.S. history performance among ethnic/racial groups showed little or no variation across levels of parental education.

The data in Table 2.12 show that the racial and ethnic differences in average U.S. history scores persist across different levels of parental educational attainment. This runs somewhat counter to previous findings from other studies. The National Education Longitudinal Study of 1988, with more complete measures of socioeconomic status, found substantial reductions in achievement differences associated with racial/ethnic group membership after accounting for family resources.[3] In addition, the College Board has found that racial differences on the Scholastic Aptitude Test are diminished somewhat when family income differences are taken into account.[4] So, the NAEP findings should be interpreted carefully in relation to these other results.

In interpreting theses findings, it is important to understand that student achievement is a result of multiple factors including educational experiences, resources from the home, and the larger social environment. These factors may also differ depending on the students' racial/ethnic groups and thus contribute — along with parents' educational level — to achievement differences. Such factors might contribute to reasonable explanations for why parents' educational levels might be associated differently with student achievement for different racial/ethnic groups.

Summary

The NAEP U.S. history assessment showed patterns of performance among reporting subgroups, some of which consistently performed better than others on the composite scale. These patterns of performance were also evident in interactions among those reporting subgroups.

- Among the different regions of the nation, students' scores varied at each grade level. Grade 8 students from the Southeast, for example, had lower scores than did grade 8 students from the other regions.
- Generally, the higher the level of parental education, the higher the level of student performance.
- At all grades, White and Asian students had higher scores than did Black and Hispanic students. The differences between White twelfth graders and their Black and Hispanic peers were significant across parental education levels.
- At grade 12, males scored higher than females. No significant differences between males and females were found at grades 4 and 8.
- Students in nonpublic schools outperformed public school students. These differences were generally significant regardless of parental education levels.

Endnotes

1. Looker, E.D. (1989). Accuracy of proxy reports of parental status characteristics. *Sociology of Education, 62*(4), 257-276.
2. As a result of the Elementary and Secondary Education Act reauthorized by Congress in 1994, the federal program formerly referred to as "Chapter One" was renamed "Title I."
3. Green, P.J., Dugone, B.L., Ingels, S.J., & Camburn, E. (1995). *A profile of the American high school senior in 1992.* Washington, DC: National Center for Education Statistics, NCES 95-384.
4. College Entrance Examination Board and Educational Testing Service (1995). *College bound seniors national profile report: SAT program test takers 1995.* Additional unpublished tables.

CHAPTER 3

U.S. History Achievement Levels

Introduction

The average U.S. history scores of our nation's students, presented in the previous chapter, can be explored further by considering the percentage of students who attained specific levels of achievement. Viewing students' performance from this perspective provides some insight into the adequacy of students' knowledge and skills and the extent to which they achieved expected levels of performance.

Three U.S. history achievement levels — *Basic*, *Proficient*, and *Advanced* — were established by the National Assessment Governing Board (NAGB) for use in reporting NAEP results. As described in Chapter 1, the *Basic* level denotes partial mastery of the knowledge and skills that are fundamental for proficient work at each grade. The *Proficient* level represents solid academic performance and competency over challenging subject matter. The *Advanced* level signifies superior performance. A more detailed description of the achievement levels is presented in Figure 3.1.

The National Education Statistics Act of 1994 requires that the National Assessment Governing Board develop "appropriate student performance levels" for reporting NAEP results. The NAEP law requires that these levels be "used on a developmental basis until the Commissioner of Education Statistics determines . . . that such levels are reasonable, valid, and informative to the public." It requires the Commissioner and the Governing Board to make clear the developmental status of such levels.

The student achievement levels in this report have been developed and adopted by the National Assessment Governing Board, NAEP's independent policymaking body, with contributions from a wide variety of educators, business and government leaders, and interested citizens. These levels of student achievement have been established to help Americans answer two questions that are important to parents and to all citizens in the communities and nation in which we live. These questions are: "What should students know and be able to do as they progress and graduate from school?" and "How good is good enough in terms of student achievement on NAEP?" Answering these questions obviously involves judgments. The NAGB is not suggesting that there is a single answer to these questions. Rather, the Board is trying to put forward reasonable judgments that can inform citizens across America — information they can use to answer these questions in their own schools and communities.

Developing carefully considered judgments about "what students should know and be able to do" and "how good is good enough" is both difficult and controversial. The Governing Board believes that these questions are so important that answers must be sought in an informed, responsible way. The process is subject to revision and refinement as appropriate.

Figure 3.1 U.S. History Achievement Levels

GRADE 4	
BASIC (195)	Fourth-grade students performing at the Basic level should be able to identify and describe a few of the most familiar people, places, events, ideas, and documents in American history. They should be able to explain the reasons for celebra ig most national holidays, have some familiarity with the geography of their own state and the United States, and be able to express in writing a few ideas about a familiar theme in American history.
PROFICIENT (243)	Fourth-grade students performing at the Proficient level should be able to identify, describe and comment on the significance of many historical people, places, ideas, events, and documents. They should interpret information from a variety of sources, including texts, maps, pictures, and timelines. They should be able to construct a simple timeline from data. These students should recognize the role of invention and technological change in history. They should also recognize the ways in which geographic and environmental factors have influenced life and work.
ADVANCED (276)	Fourth-grade students performing at the Advanced level should have a beginning understanding of the relationships between people, places, ideas, events and documents. They should know where to look for information, including reference books, maps, local museums, interviews with family and neighbors, and other sources. They should be able to use historical themes to organize and interpret historical topics, and to incorporate insights from beyond the classroom into their understanding of history. These students should understand and explain the role of invention and technological change in history. They should also understand and explain the ways in which geographic and environmental factors have influenced life and work.
GRADE 8	
BASIC (252)	Eighth-grade students performing at the Basic level should be able to identify and place in context a range of historical people, places, events, ideas, and documents. They should be able to distinguish between primary and secondary sources. They should have a beginning understanding of the diversity of the American people and the ways in which people from a wide variety of national and cultural heritages have become part of a single nation. Eighth-grade students at the basic level should also have a beginning understanding of the fundamental political ideas and institutions of American life and their historical origins. They should be able to explain the significance of some major historical events.
PROFICIENT (294)	Eighth-grade students performing at the Proficient level should be able to explain the significance of people, places, events, ideas, and documents, and to recog ize the connection between people and events within historical contexts. They should understand and be able to plain the opportunities, perspectives and challenges associated with a diverse cultural population. They should incorporate geographic, technological, and other considerations in their understanding of events and should have knowledge of significant political ideas and institutions. They should be able to communicate ideas about historical themes while citing evidence from primary and secondary sources to support their conclusions.
ADVANCED (327)	Eighth-grade students performing at the Advanced level should recognize significant themes and movements in history and begin to understand particular events in light of these themes and movements. They should have an awareness of continuity and change over time and be able to draw relevant analogies between past events and present-day situations. They should be able to frame questions about historical topics and use multiple sources to develop historical generalizations and interpretations. They should be able to explain the importance of historical themes, including some awareness of their political, social, and economic dimensions.
GRADE 12	
BASIC (294)	Twelfth-grade students performing at the Basic level should be able to identify the significance of many people, places, events, dates, ideas, and documents in U.S. history. They should also recognize the importance of unity and diversity in the social and cultural history of the United States, and an awareness of America's changing relationships with the rest of the world. They should have a sense of continuity and change in history and be able to relate relevant experience from the past to their understanding of contemporary issues. They should recognize that history is subject to interpretation and should understand the role of evidence in making an historical argument.
PROFICIENT (325)	Twelfth-grade students performing at the Proficient level should understand particular people, places, events, ideas, and documents in historical context, with some awareness of the political, economic, geographic, social, religious, technological, and ideological factors that shape historical settings. They should be able to communicate reasoned interpretations of past events, using historical evidence effectively to support their positions. Their written arguments should reflect some in-depth grasp of issues and refer to both primary and secondary sources.
ADVANCED (355)	Twelfth-grade students achieving at the Advanced level should demonstrate a comprehensive understanding of events and sources of U.S. history. Recognizing that history is subject to interpretation, they should be able to evaluate historical claims critically in light of the evidence. They should understand that important issues and themes have been addressed differently at different times and that America's political, social, and cultural traditions have changed over time. They should be able to write well-reasoned arguments on complex historical topics and draw upon a wide range of sources to inform their conclusions.

The student achievement levels in this report, approved by the Governing Board, are the result of countless hours of work. The levels are based on preliminary descriptions developed as part of the national consensus process to determine the assessment design and content. The Board's contractor, American College Testing (ACT), which has extensive experience in standard setting in many fields, designed the achievement level-setting process. This process was reviewed by scores of individuals, including policymakers, professional organizations, teachers, parents, and other members of the general public. To develop the levels, ACT convened a cross section of educators and interested citizens from across the nation and asked them to recommend what students should know and be able to do in U.S. history. Prior to adopting these levels of student achievement, the Board engaged a large number of persons to comment on the recommended levels and to review the results.

The result of the achievement level-setting process is a set of achievement level descriptions, a set of achievement level cut points on the 500-point NAEP scale, and exemplar questions. The cut points are minimum scores that define *Basic*, *Proficient*, and *Advanced* performance at grades 4, 8, and 12. At present, evaluations conducted on the level-setting process and critiques of those evaluations have provided mixed reviews. Therefore, both the Governing Board and the Commissioner of Education Statistics regard the achievement levels as developmental; they should not be interpreted as statistically conclusive. Because these levels are still considered developmental, the reader of tnis report is advised to consider that status when interpreting the results. The reader should recognize that the results are based on the judgments of panels, approved by the Governing Board, of what *Basic*, *Proficient*, and *Advanced* students should know and be able to do in each subject assessed, as well as on their judgments regarding what percent of students at the borderline for each level should answer each test question correctly. The latter information is used in translating the achievement level descriptions into cut points on the NAEP scale. NCES uses these levels in reporting NAEP results, but it does not currently adjudicate the reliability or validity of these achievement levels. Rather, they are reported directly as adopted by the Governing Board.

The U.S. history results presented here for grades 4, 8, and 12 illustrate one of the difficulties of setting achievement levels. The Governing Board is concerned about the discrepancy between actual student performance described in this report and the expectations for performance that are contained in the achievement levels. Simply stated, students are not performing as well on the NAEP U.S. history assessment, particularly at grade 12, as the Governing Board and the many panelists and reviewers think these students should perform. For example, most students take at least one high school course in U.S. history by the end of the eleventh grade. Yet the achievement levels indicate that more than half (57 percent) of twelfth graders are performing below the *Basic* level, with 1 percent scoring at the *Advanced* level. In contrast, data from the College Board show that about 2.4 percent of all graduating seniors score well enough on the Advanced Placement exam in U.S. history to be considered qualified for college credit.

Since NAEP is a cross-sectional survey of student achievement, it cannot readily identify cause and effect relationships to explain why students scored high or low. Although one hypothesis is that students' performance was found to be too low because the achievement levels are set too high, NAGB does not believe that this is the case. At present, studies on these achievement levels, conducted by ACT, have pointed in opposite directions — one suggested the levels were too high, the other that they were too low. The Governing Board intends to look carefully at this gap between expected and actual performance, and encourages others to do so as well.

Nevertheless, there are several other hypotheses that might account for this gap between actual student scores and the achievement levels. Motivation, particularly at grade 12, is a perennial question in an assessment like NAEP for which there are no stakes or rewards for students to do well. (However, it is not clear why students should be less motivated in taking this history assessment than other NAEP assessments in which higher percentages of students reached the various cutpoints.) There may be differences between what is taught in the broad array of U.S. history classes and the content of this NAEP assessment. A lack of consistency between the grade levels at which the subject is taught and the NAEP assessment of grades 4, 8, and 12 could account for some of this discrepancy. The judges for the twelfth-grade levels may have had relatively higher expectations than judges for the other grades. Finally, the difference between more conventional testing practices in some classrooms and the NAEP assessment questions may be another factor. NAEP includes a variety of questions, from multiple-choice questions to constructed-response tasks that require students to apply knowledge and demonstrate skills by writing their answers.

Many of these factors, or a combination of all of them, could explain the gap between standards for student performance contained in the NAGB achievement levels and the actual performance on the NAEP 1994 U.S. history assessment. The National Assessment Governing Board urges all who are concerned about "what students should know and be able to do" and "how good is good enough" to read and interpret these performance levels recognizing that this is a developing, judgmental process and is subject to various interpretations. The decision to include the levels in NAEP reports is an attempt to make the assessment results more useful for parents, educators, and policymakers by providing performance standards against which to measure educational progress.

The U.S. history achievement levels attained by fourth-, eighth-, and twelfth-grade students in the NAEP assessment are presented in this chapter. Results are displayed for the nation, regions of the country, and major reporting subgroups. The same cautions stated in Chapter 2 are warranted when interpreting performance differences among subgroups.

U.S. History Achievement Levels for the Nation

The percentages of fourth-, eighth-, and twelfth-grade students who performed at or above the three U.S. history achievement levels are shown in Table 3.1. This table provides the results for the nation as a whole and by region. The majority of students at grade 4 (64 percent) and grade 8 (61 percent) performed at or above the *Basic* achievement level. However, at grade 12, more than half of the students (57 percent) performed below the *Basic* level. The *Proficient* level, which is described as representing solid academic performance, was reached by 17 percent of fourth graders, 14 percent of eighth graders, and 11 percent of twelfth graders. Few students at any grade were at the *Advanced* level: 2 percent at grade 4, and 1 percent at grades 8 and 12.

Overall, few of the achievement level differences in performance by region are statistically significant. No significant differences were observed at the fourth grade or among students reaching the *Advanced* level at any grade.

The largest number of significant differences was observed among eighth-grade students performing at or above the *Basic* and *Proficient* levels. Higher percentages of eighth graders in the Central and Northeast regions reached the *Basic* and *Proficient* levels than in the Southeast or West. The percentage of grade 8 students in the West at or above the *Basic* level was also higher than that in the Southeast.

At the twelfth grade, the only significant difference was that the percentages of students at or above *Basic* in both the Northeast and Central regions were higher than the percentage in the Southeast.

This performance in terms of achievement levels represents a lower level of attainment than has been observed in other NAEP surveys. This discrepancy is greatest at grade 12. For example, on the 1994 geography assessment, 70 percent of students scored at or above the *Basic* level, 27 percent at or above *Proficient*, and 2 percent at *Advanced*. This compares with U.S. history results of 43 percent at or above *Basic*, 11 percent at or above *Proficient*, and 1 percent at *Advanced*.

These differences at the *Proficient* level mean that students were less likely to exhibit what independent panels of judges viewed as solid grade-level performance in U.S. history than they did in geography and other subjects. This difference may be the result of actual weakness in student achievement in U.S. history. However, when interpreting cross-subject differences readers should note that the achievement levels for each subject are set independently of work in other subjects, and that no attempt is made to make the levels statistically similar across subjects. Judges charged with setting achievement levels do work with common policy definitions (see page 6). However, the subject-specific definitions vary. Finally, judges in different subjects may bring different assumptions, expectations, and standards to the process of setting levels, and these may lead to variations across subjects.

TABLE 3.1

U.S. History Achievement Levels for the Nation and by Region Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

		Percentage of Students			
	Percentage of All Students	At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Region					
Northeast	22 (0.7)	3 (0.7)	18 (2.1)	63 (2.4)	37 (2.4)
Southeast	23 (1.0)	2 (0.5)	15 (1.5)	61 (2.3)	39 (2.3)
Central	25 (0.8)	3 (1.0)	20 (2.6)	71 (2.8)	29 (2.8)
West	30 (0.6)	1 (0.8)	16 (1.8)	61 (2.5)	39 (2.5)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Region					
Northeast	20 (0.8)	1 (0.3)	19 (1.7)	69 (2.0)	31 (2.0)
Southeast	25 (0.9)	0 (0.2)	9 (0.8)	51 (1.9)	49 (1.9)
Central	24 (0.6)	1 (0.4)	17 (1.2)	69 (2.3)	31 (2.3)
West	31 (0.8)	1 (0.2)	11 (1.2)	58 (1.3)	42 (1.3)
Grade 12					
Nation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
Region					
Northeast	20 (0.5)	1 (0.4)	13 (1.5)	46 (2.4)	54 (2.4)
Southeast	23 (0.8)	0 (0.3)	8 (1.4)	37 (1.9)	63 (1.9)
Central	27 (0.7)	1 (0.5)	11 (1.2)	45 (2.0)	55 (2.0)
West	30 (0.7)	1 (0.2)	10 (1.2)	43 (2.3)	57 (2.3)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

The estimates of population percentages reported as zero (and standard errors reported as 0.0) are actually nonzero, but rounded to zero when reporting to the nearest integer (or nearest tenth in the case of the standard errors).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

U.S. History Achievement Levels for Major Reporting Subgroups

In this section, variations in performance among the major reporting subgroups are discussed. Again, the discussion is confined to those differences that are statistically significant.

Race/Ethnicity. The percentages of students in each group who performed at or above each achievement level are presented in Table 3.2. Statistically significant differences in performance among racial/ethnic groups were found more often at grades 4 and 8 than at grade 12.

At grades 4 and 8, the percentage of White students reaching the *Advanced* level was higher than the percentages of Black or Hispanic students, though it is

TABLE 3.2

U.S. History Achievement Levels by Race/Ethnicity Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage of All Students	Percentage of Students			
		At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Race/Ethnicity					
White	69 (0.3)	3 (0.5)	22 (1.4)	74 (1.1)	26 (1.1)
Black	15 (0.1)	0 (0.1)	4 (1.0)	36 (1.8)	64 (1.8)
Hispanic	11 (0.2)	1 (0.4)	6 (1.2)	41 (3.6)	59 (3.6)
Asian	2 (0.2)	4 (2.6)	22 (5.1)	64 (4.7)	36 (4.7)
Pacific Islander	1 (0.2)	3 (2.9)	16 (4.7)	59 (7.1)	41 (7.1)
American Indian	2 (0.3)	0 (1.1)	9 (2.7)	51 (6.9)	49 (6.9)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Race/Ethnicity					
White	69 (0.2)	1 (0.2)	17 (0.8)	71 (1.1)	29 (1.1)
Black	15 (0.1)	0 (0.1)	4 (0.6)	33 (2.3)	67 (2.3)
Hispanic	11 (0.1)	0 (0.1)	5 (0.7)	41 (2.3)	59 (2.3)
Asian	2 (0.1)	2 (1.2)	23 (4.2)	72 (5.1)	28 (5.1)
Pacific Islander	1 (0.3) !	1 (0.8) !	11 (4.4) !	52 (8.7) !	48 (8.7) !
American Indian	1 (0.3) !	0 (0.0) !	5 (2.6) !	42 (5.8) !	58 (5.8) !
Grade 12					
ation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
ace/Ethnicity					
White	74 (0.4)	1 (0.2)	13 (0.8)	50 (1.2)	50 (1.2)
Black	12 (0.2)	0 (0.1)	2 (0.8)	17 (1.6)	83 (1.6)
Hispanic	9 (0.3)	0 (0.3)	4 (0.7)	22 (2.1)	78 (2.1)
Asian	3 (0.2)	2 (1.1)	16 (4.6)	46 (5.4)	54 (5.4)
Pacific Islander	1 (0.2)	1 (0.6)	7 (3.8)	33 (6.4)	67 (6.4)
American Indian	1 (0.2) !	0 (0.0) !	5 (2.3) !	30 (7.6) !	70 (7.6) !

ferences between the groups may be partially explained by other factors not included in this table

standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the mate for the sample.

centages of students in the subgroups may not total 100 due to rounding or, in the case of the race/ethnicity variable, because some students categorized themselves as "other."

estimates of population percentages reported as zero are actually nonzero, but round to zero when reporting to the nearest integer.

terpret with caution any comparisons involving this statistic. The nature of the sample does not allow for accurate determination of the variability of this value

IRCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

important to note that only 1 or 2 percent of the total population at these grades reached this level. Also, the percentages of White students at or above the *Proficient* level at both grades were higher than those of Black or Hispanic students, and higher than those of American Indian students at grade 4. At both grades, the percentages of Asian students were higher than those of Black or Hispanic students who reached that level.

The largest number of differences was evident at or above the *Basic* level, although the pattern was similar to that shown at the other two levels. At both grades 4 and 8, the percentages of White and Asian students at or above the *Basic* level were higher than the percentages of Black or Hispanic students. In addition, at the fourth grade, a higher percentage of White students than American Indian students, and a higher percentage of Pacific Islander students than Black students, were at or above the *Basic* level.

At grade 12 there were fewer differences, but they followed a similar pattern. The percentage of White students at the *Advanced* level was higher than the percentage of Black students. (Again, however, only 1 percent of the total population reached this level, so the practical significance of this difference is slight.) The percentages of White students at or above the *Basic* and *Proficient* levels were higher than the percentages of Black or Hispanic students. Finally, the percentage of Asian twelfth graders at or above the *Basic* level was higher than the percentages of their Black and Hispanic counterparts.

As mentioned in Chapter 2, the natures of the grades 8 and 12 American Indian student samples and the grade 8 Pacific Islander student sample do not allow accurate determination of standard errors. For this reason, differences among those samples and other racial/ethnic groups are not discussed.

Gender. The percentages of males and females reaching each of the achievement levels are presented in Table 3.3. Two significant differences were evident: the percentages of twelfth-grade males at or above the *Proficient* and *Basic* levels were higher than the percentages of females.

TABLE 3.3

U.S. History Achievement Levels by Gender Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

		Percentage of Students			
	Percentage of All Students	At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Gender					
Male	50 (0.8)	2 (0.4)	18 (1.4)	62 (1.6)	38 (1.6)
Female	50 (0.8)	2 (0.4)	16 (1.1)	65 (1.4)	35 (1.4)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Gender					
Male	50 (0.5)	1 (0.3)	15 (0.8)	61 (1.0)	39 (1.0)
Female	50 (0.5)	1 (0.1)	13 (0.8)	61 (1.3)	39 (1.3)
Grade 12					
Nation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
Gender					
Male	50 (0.8)	1 (0.3)	12 (0.7)	45 (1.2)	55 (1.2)
Female	50 (0.8)	1 (0.2)	9 (0.8)	40 (1.4)	60 (1.4)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Parents' Level of Education. As shown in Table 3.4, parental education and student achievement are positively related. This finding mirrors the average scale score results discussed in the previous chapter. At grades 4 and 8, the percentages of students who reported that at least one parent had graduated from college and who reached the *Advanced* level were higher than the percentages of their counterparts who reported that their parents graduated from high school. At grade 12, the percentage of students who reported that at least one parent graduated from college and who reached the *Advanced* level was higher than that of twelfth graders who reported that their parents had a high school or lower level of education.

At the *Proficient* and *Basic* levels, the patterns were similar. At all grades, the percentages of students at or above both the *Basic* and *Proficient* levels who reported that their parents had achieved a given level of education were higher than those who reported lower levels of parental education. For example, at grade 8, the percentage of students who reported that at least one parent graduated from college and who were at or above the *Proficient* level was higher than that of students who reported that at least one parent had achieved some

TABLE 3.4

U.S. History Achievement Levels by Parents' Highest Education Level Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

		Percentage of Students			
	Percentage of All Students	At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Parents' Education Level					
Graduated College	43 (0.9)	4 (0.6)	25 (1.7)	74 (1.2)	26 (1.2)
Some Education After High School	8 (0.4)	3 (1.0)	21 (2.6)	74 (2.8)	26 (2.8)
Graduated High School	12 (0.5)	1 (0.4)	10 (1.3)	57 (2.7)	43 (2.7)
Did Not Finish High School	4 (0.3)	0 (0.0)	2 (1.2)	37 (4.1)	63 (4.1)
I Don't Know	34 (0.8)	0 (0.3)	10 (0.9)	55 (1.7)	45 (1.7)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Parents' Education Level					
Graduated College	42 (1.0)	1 (0.3)	22 (1.1)	74 (1.2)	26 (1.2)
Some Education After High School	19 (0.5)	0 (0.2)	14 (1.1)	68 (1.3)	32 (1.3)
Graduated High School	23 (0.8)	0 (0.1)	7 (0.9)	50 (1.4)	50 (1.4)
Did Not Finish High School	7 (0.4)	0 (0.0)	3 (0.7)	37 (2.2)	63 (2.2)
I Don't Know	9 (0.4)	0 (0.1)	3 (1.1)	36 (2.2)	64 (2.2)
Grade 12					
Nation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
Parents' Education Level					
Graduated College	45 (1.0)	1 (0.4)	17 (1.0)	56 (1.3)	44 (1.3)
Some Education After High School	25 (0.7)	1 (0.3)	8 (1.1)	42 (1.9)	58 (1.9)
Graduated High School	20 (0.7)	0 (0.1)	4 (0.8)	29 (1.5)	71 (1.5)
Did Not Finish High School	7 (0.4)	0 (0.2)	1 (0.6)	15 (1.9)	85 (1.9)
I Don't Know	3 (0.2)	0 (0.2)	1 (0.9)	12 (3.5)	88 (3.5)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

The estimates of population percentages reported as zero are actually nonzero, but round to zero when reporting to the nearest integer

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

education after high school. The latter group had a higher percentage of students at or above the *Proficient* level than that of students who reported tha at least one parent graduated from high school. These students, in turn, showed higher achievement than those who reported that neither parent had finished high school. The exception to this rule is at grade 4, where there were no significant differences in the achievement level attainments of children who reported that at least one parent graduated from college and those who reported that at least one parent had some education after high school.

As shown in Table 3.4, one-third of fourth graders and nearly one-tenth of eighth graders did not know their parents' level of education. Also, the accuracy of student self-reported data may be open to some question, as previously noted. Nonetheless, the positive relationship between parental education and achievement in U.S. history remains striking.

Type of Location. Achievement level results by type of location are shown in Table 3.5. Type of location categories indicate the geographic locations of students' schools and are not intended to indicate or imply social or economic meanings for these location types. At the fourth grade, the only statistically significant differences were observed at the *Basic* level. The percentage of students attending schools in urban fringe/large town locations who were at or above the *Basic* level was higher than those of students in the other two locations. At grade 8, the only significant difference was that the percentage of students at or above the *Basic* level attending schools in urban fringe/large-town locations was higher than that of their central city counterparts. At grade 12, a pattern was evident at both the *Proficient* and the *Basic* levels: the percentages of students attending schools in rural/small town locations who were at or above both the *Basic* or *Proficient* levels were lower than those of students in the other two categories.

TABLE 3.5

U.S. History Achievement Levels by Type of Location Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

		Percentage of Students			
	Percentage of All Students	At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Type of Location					
Central City	35 (2.1)	2 (0.6)	15 (1.4)	57 (2.2)	43 (2.2)
Urban Fringe/Large Town	43 (2.4)	3 (0.6)	20 (1.5)	70 (1.7)	30 (1.7)
Rural/Small Town	22 (2.1)	1 (0.7)	15 (2.3)	62 (2.7)	38 (2.7)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Type of Location					
Central City	36 (2.4)	1 (0.3)	14 (1.0)	57 (1.9)	43 (1.9)
Urban Fringe/Large Town	38 (2.9)	1 (0.2)	15 (1.1)	65 (1.7)	35 (1.7)
Rural/Small Town	26 (1.8)	0 (0.2)	11 (1.0)	61 (2.3)	39 (2.3)
Grade 12					
Nation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
Type of Location					
Central City	31 (2.2)	1 (0.4)	11 (1.1)	43 (1.8)	57 (1.8)
Urban Fringe/Large Town	43 (2.7)	1 (0.2)	12 (1.0)	47 (1.5)	53 (1.5)
Rural/Small Town	26 (2.0)	0 (0.3)	7 (0.8)	36 (1.7)	64 (1.7)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

The estimates of population percentages reported as zero (and standard errors reported as 0.0) are actually nonzero, but rounded to zero when reporting to the nearest integer (or nearest tenth in the case of the standard errors).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Title I Participation. Table 3.6 presents the achievement levels for each grade by students' participation in Title I programs. At grade 4, the percentages of students who do not receive Title I services who were at each achievement level were higher than those of students who do receive Title I services. At grade 8, the differences between the two groups were significant for the percentages of students at or above the *Basic* and *Proficient* levels. Grade 12 differences are not discussed here because the nature of the grade 12 sample does not allow for accurate estimation of the variability of the percentages for Title I recipients.

TABLE 3.6

U.S. History Achievement Levels by Title I Participation Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

		Percentage of Students			
	Percentage of All Students	At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Title I Participation					
Yes	14 (1.3)	0 (0.2)	1 (0.4)	22 (2.6)	78 (2.6)
No	86 (1.3)	2 (0.4)	20 (1.1)	71 (1.2)	29 (1.2)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Title I Participation					
Yes	7 (1.0)	0 (0.0)	1 (0.6)	24 (3.1)	76 (3.1)
No	93 (1.0)	1 (0.2)	14 (0.7)	64 (1.0)	36 (1.0)
Grade 12					
Nation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
Title I Participation					
Yes	2 (0.5) !	0 (0.3) !	1 (0.5) !	9 (2.4) !	91 (2.4) !
No	98 (0.5)	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

The estimates of population percentages reported as zero are actually nonzero, but round to zero when reporting to the nearest integer.

! Interpret with caution any comparisons involving this statistic. The nature of the sample does not allow for accurate determination of the variability of this value

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Type of School. Results for nonpublic and public school students are presented in Table 3.7. At all three grades, the percentages of nonpublic school students at or above both the *Basic* and *Proficient* levels were higher than the percentages of students attending public schools. At the eighth grade, nonpublic school students were also more likely to reach the *Advanced* level than were public school students.

As noted earlier, the reader is cautioned against using these data to make simplistic inferences about the relative effectiveness of public and nonpublic schools. Differences between the two types of schools are in part related to socioeconomic factors and sociological factors, such as levels of parental involvement.

TABLE 3.7

U.S. History Achievement Levels by Type of School Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

		Percentage of Students			
	Percentage of All Students	At Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4					
Nation	100	2 (0.3)	17 (1.0)	64 (1.1)	36 (1.1)
Type of School					
Public Schools Only	90 (0.8)	2 (0.3)	16 (1.1)	62 (1.2)	38 (1.2)
Nonpublic Schools Only	10 (0.8)	3 (0.7)	26 (1.9)	82 (2.1)	18 (2.1)
Catholic Schools	6 (0.7)	2 (0.7)	24 (2.3)	81 (2.6)	19 (2.6)
Other Nonpublic Schools	4 (0.5)	5 (1.4)	29 (3.9)	83 (4.0)	17 (4.0)
Grade 8					
Nation	100	1 (0.1)	14 (0.6)	61 (0.9)	39 (0.9)
Type of School					
Public Schools Only	90 (0.9)	1 (0.1)	12 (0.6)	59 (1.0)	41 (1.0)
Nonpublic Schools Only	10 (0.9)	2 (0.6)	28 (1.8)	84 (1.2)	16 (1.2)
Catholic Schools	6 (0.6)	2 (0.8)	29 (2.3)	85 (1.7)	15 (1.7)
Other Nonpublic Schools	4 (0.6)	2 (0.8)	26 (3.4)	83 (2.8)	17 (2.8)
Grade 12					
Nation	100	1 (0.2)	11 (0.7)	43 (1.1)	57 (1.1)
Type of School					
Public Schools Only	89 (1.1)	1 (0.2)	10 (0.7)	41 (1.2)	59 (1.2)
Nonpublic Schools Only	11 (1.1)	1 (0.4)	18 (1.3)	59 (2.2)	41 (2.2)
Catholic Schools	6 (0.9)	1 (0.4)	18 (2.0)	57 (3.8)	43 (3.8)
Other Nonpublic Schools	5 (0.6)	2 (0.8)	19 (2.7)	61 (2.9)	39 (2.9)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Summary

The performance of students on the NAEP 1994 U.S. history assessment reflected the ambitious nature of the framework that guided the development of the assessment, and the challenging expectations based on the achievement levels. For the most part, performances of reporting subgroups reflected patterns of average scale scores noted in Chapter 2 of this report.

- For the nation as a whole, few students reached the *Proficient* level.
- Generally, higher percentages of White and Asian students reached the *Basic* and *Proficient* levels than did Black and Hispanic students.
- At grade 12, the percentages of males at or above the *Basic* and *Proficient* levels were higher than those of females. (There were no significant differences at grades 4 and 8 between males and females.)
- Generally, the higher the level of parental education, the greater the percentage of students reaching a given achievement level.
- Higher percentages of nonpublic school students reached the *Basic* and *Proficient* levels than did percentages of public school students.

CHAPTER 4

Contexts in Which Students Learn History

Classroom instruction is at the core of history education, but history learning can take place in a variety of settings both inside and outside the classroom. Many teachers use trips to historical sites or museums, multimedia presentations, and a variety of teaching strategies to enhance history learning. Outside of school, students are exposed to history in many forms. Children may learn history from television, film documentaries or dramas, extracurricular books, newspapers and magazines, and from conversations with family members or other adults. Family and regional history are often learned informally. Previous chapters of this report have examined what students know and can do in history; if we wish to gain an understanding of why they know what they do, we must examine some of the various contexts in which history learning occurs.

The Extent of Students' Social Studies and U.S. History Instruction

As part of the NAEP 1994 U.S. history assessment, students were asked a series of questions concerning the nature of the social studies or history instruction they had received. In addition, information concerning social studies or history instruction was collected from the teachers of fourth- and eighth-grade students participating in the assessment. The results illustrate the nature of U.S. history instruction in the nation's schools.

Social Studies and History Instruction at Grade 4. The *NAEP 1994 U.S. History Framework* that served as the blueprint for the assessment recognized that most fourth graders do not have a formal class in U.S. history. (Attention was paid in the framework to ensuring coverage of material that is likely to be addressed in fourth-grade classrooms, such as state history.) However, the majority of fourth graders did report having regular social studies or history classes.

Approximately two-thirds (69 percent) of fourth graders reported having a social studies or history class at least three times a week. Only 7 percent reported never or hardly ever having had such classes. Table 4.1 presents the percentage of students and average scale scores according to the frequency with which they have taken social studies or history classes.

TABLE 4.1 **Students' Reports on Social Studies or History Course-Taking Grade 4**

THE NATION'S REPORT CARD 1994

	Percentage	Average Scale Score
How often do you usually have social studies or history class in school?		
Almost Every Day	44 (1.9)	209 (1.3)
3 to 4 Times a Week	25 (1.1)	213 (1.4)
1 to 2 Times a Week	19 (1.1)	200 (1.8)
Less than 1 Time a Week	6 (0.5)	184 (4.0)
Never or Hardly Ever	7 (0.6)	182 (3.4)
Have you ever studied the history of the United States in school?		
Yes	79 (0.8)	208 (1.0)
No	21 (0.8)	194 (1.9)

Differences between the groups may be partially explained by other factors not included in this table

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Students who reported having social studies or history class every day or three to four times a week had higher average scores than did those reporting having such classes less frequently. In addition, students who reported having classes once or twice a week outperformed, on average, those who reported having such classes less than weekly or never or hardly ever.

When specifically asked if they have ever studied U.S. history, the majority of fourth graders, 79 percent, reported that they had (see Table 4.1). Those students who reported studying U.S. history had a significantly higher average score than did those who reported never having studied the subject.

When the teachers of the fourth graders were asked about the focus of their social studies classes, it was evident that few students were currently taking social studies classes that focused primarily on U.S. history. Approximately half of the students had teachers whose classes focused on either social studies (25 percent) or social studies integrated with other subjects (24 percent). Another third (36 percent) had classes that focused on state history. Only 7 percent had classes for which the teacher reported that the focus was U.S. history. Table 4.2 presents the teachers' reports and the students' average scale scores.

TABLE 4.2 **Teachers' Reports on Focus of Social Studies/ History Teaching Grade 4**

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage	Average Scale Score
State History	36 (2.8)	207 (2.0)
United States History	7 (1.6)	209 (6.7)
Social Studies	25 (2.8)	205 (2.8)
Social Studies Integrated with Other Disciplines	24 (2.7)	198 (2.2)
Geography	6 (1.2)	212 (4.5)
Other	2 (0.7)	225 (8.1)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Except for the "other" category, no significant differences in average scores were evident among students of teachers reporting different foci. (The average score for "other" was significantly higher than that for the "social studies integrated with other disciplines" focus.)

Examined together, the results of the NAEP questions concerning fourth-grade students' in-school exposure to U.S. history indicated that

- the majority of fourth graders reported taking social studies or history classes, and
- over three-quarters of fourth graders also said they have had some exposure to U.S. history, but
- according to their teachers, most students are not currently exposed to social studies/history classes that focus on U.S. history.

Frequent social studies or history classes and some exposure to U.S. history in the classroom are positively related to scores on the NAEP U.S. history assessment, even though fourth graders are not currently involved in focused study of the subject.

U.S. History Course-Taking at Grades 8 and 12. Students in grades 8 and 12 were asked whether they were currently taking a U.S. history class or if they had taken such a class previously. Even though NAEP assessed U.S. history at grades 8 and 12, background questionnaire responses indicate that the subject is typically taught in grades 8 and 11. When eighth and twelfth graders were asked if they were currently taking a U.S. history course, the majority of eighth graders (83 percent) said they were, while only a third of twelfth graders (34 percent) said they were taking such a course.

Table 4.3 presents the percentage and average scale scores of eighth graders currently taking a U.S. history course. The table also presents the results for twelfth graders who reported they are currently taking a U.S. history course and results for those who reported they took such a course in grade 11.

TABLE 4.3 **Students' Reports on U.S. History Course-Taking Grades 8 and 12**

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage	Average Scale Score
Grade 8		
Currently Taking a U.S. History Course		
Yes	83 (1.6)	261 (0.7)
No	17 (1.6)	253 (1.6)
Grade 12		
Currently Taking a U.S. History Course		
Yes	34 (1.6)	283 (1.3)
No	66 (1.6)	288 (0.9)
Took a U.S. History Course in Grade 11		
Yes	79 (1.5)	288 (0.8)
No	18 (1.5)	287 (1.6)
I Don't Know	3 (0.3)	261 (3.2)

Differences between the groups may be partially explained by other factors not included in this table

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

At grade 8, the average score of students currently taking a U.S. history class was significantly higher than that of students not taking such a class. At the twelfth grade, the reverse is true — students currently taking a U.S. history course had significantly lower performance, on average, than those not taking such a course. However, when twelfth graders who reported taking a U.S. history course in eleventh grade are compared to those reporting that they did not, no significant differences in average scale scores are seen.

At the middle and high school levels, the relationship between course-taking behaviors and scale scores varies. In grade 8, where the assessment coincides with the majority of students taking a class in the subject, a positive relationship between course-taking and performance is seen. At grade 12, where the majority of students are not taking a U.S. history class, a negative relationship is seen. Given that the twelfth graders taking U.S. history classes in their senior year are not following the course-taking pattern of the majority of their peers, it is possible that students taking U.S. history in grade 12 have different characteristics from the majority of students.

TABLE 4.4

Students' Reports on Number of Semesters of History, Geography, or Social Studies Taken Grade 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage	Average Scale Score
None	0 (0.1)	***
1 or 2	7 (0.5)	263 (2.1)
3 or 4	28 (0.7)	278 (1.2)
5 or 6	30 (0.9)	295 (1.1)
7 or More	35 (1.3)	300 (1.0)

Differences between the groups may be partially explained by other factors not included in this table.

Students were asked to indicate how many semesters of course work they had taken from the beginning of the 9th grade through the end of the current school year.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

The estimates of population percentages reported as zero are actually nonzero, but round to zero when reporting to the nearest integer.

*** Sample size insufficient to permit a reliable estimate (see Appendix A).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

A closer look at the amount of potential history-related course work completed by twelfth graders presents more predictable results. Table 4.4 shows that as the number of semesters of history, geography, or social studies taken increased, so did students' average scale scores.

The Content of Social Studies and History Instruction at Grades 4 and 8

As was discussed in the previous section, the majority of fourth graders have frequent social studies classes, although the classes, for the most part, do not focus on U.S. history. The majority of eighth graders, over 80 percent, reported they were currently taking a U.S. history class. In order to better understand students' experiences in these classes, this section provides a closer look at the instructional practices and materials used for history learning. (Given the small percentage of twelfth graders taking a U.S. history class — approximately one-third — the results on instructional practices and materials for grade 12 are minimal and are not presented.) All fourth- and eighth-grade students taking the NAEP U.S. history assessment and their teachers were asked about activities and materials used as part of studying history or social studies in school.

Instructional Materials

Many teachers use a wide variety of materials in their social studies and history classes. Studies have shown that incorporating a variety of teaching methods and materials can help to engage students with different learning styles. Furthermore, the richness of history as a subject can best be conveyed through the inclusion of a variety of materials.[1] In addition to standard textbooks, teachers utilize other reading materials (such as primary documents, biographies, or historical fiction), maps and globes, movies, or computers as part of their daily instruction.

As presented in the following discussion, some discrepancies exist between student — and teacher — reported frequencies for some of the instructional materials and practices examined. It is not possible to offer conclusive reasons for these discrepancies or to determine which report most accurately reflects fourth- and eighth-grade classroom activities. The reports presented represent students' and teachers' impressions of the frequency of various activities in their classrooms.

Instructional materials and strategies may be chosen for a variety of purposes. Consequently, the relationship between frequency of use of various materials and practices and scores on the NAEP U.S. history assessment cannot be interpreted in a causal fashion. The design of the assessment does not allow for the evaluation of the effectiveness of different strategies. The initial ability of the students, the particular topics being taught, and the complexity of the subject matter all need to be controlled before cause-and-effect statements can be made, and such experimental methods were not possible in a large-scale assessment.

Use of Textbooks and Other Written Material. According to both teachers and students, the use of textbooks and other written material such as historical fiction and biographies is common at both the fourth and eighth grades.

Teachers of fourth and eighth graders were asked questions about the frequency of use of textbooks, other materials such as biographies and historical fiction, and primary documents. While there are no significant relationships between student average scale scores and teacher-reported frequency of use of these materials, the frequencies themselves are informative. As is evident from Table 4.5, teachers of over 80 percent of students at grades 4 and 8 reported using textbooks at least once or twice a week. Teachers of slightly more than one-third of fourth and eighth graders reported using extra written materials, such as biographies or historical fiction, at least once a week. Over 60 percent of fourth- and eighth-grade students were taught by teachers who reported using such extra materials once or twice a month or less. Further, only 9 percent of fourth graders, and 22 percent of eighth graders, were taught by teachers who reported using primary documents at least once a week. Teachers of 62 percent of fourth graders reported never or hardly ever using primary documents, while teachers of 23 percent of eighth graders reported never or hardly ever using such documents.

As presented in Table 4.6, 71 percent of fourth graders and 84 percent of eighth graders reported reading textbooks at least once a week as part of social studies or history classes. Over a third reported reading extra written material at least once a week. Primary document use was less common; 27 percent of fourth graders and 19 percent of eighth graders reported using such documents at least once a week.

TABLE 4.5

Teachers' Reports on Use of Textbooks, Extra Written Materials, and Primary Documents Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Use of Textbooks		
Almost Every Day	43 (2.6) 207 (2.2)	45 (3.6) 259 (1.3)
Once or Twice a Week	44 (2.6) 204 (1.3)	42 (3.2) 259 (1.5)
Once or Twice a Month	8 (1.4) 204 (4.6)	8 (1.4) 266 (2.1)
Never or Hardly Ever	5 (0.9) 204 (5.0)	5 (1.8) 265 (3.6)
Use of Extra Written Materials		
Almost Every Day	6 (1.2) 208 (4.4)	3 (0.6) 254 (4.5)
Once or Twice a Week	33 (2.2) 205 (1.9)	32 (2.8) 258 (1.5)
Once or Twice a Month	46 (2.5) 204 (1.9)	47 (2.4) 263 (1.2)
Never or Hardly Ever	15 (1.7) 208 (3.2)	17 (2.2) 258 (2.5)
Use of Primary Documents		
Almost Every Day	1 (0.4) ***	2 (0.7) 268 (9.4)
Once or Twice a Week	8 (1.2) 201 (4.5)	20 (2.2) 260 (2.0)
Once or Twice a Month	29 (1.9) 208 (2.1)	55 (2.3) 261 (1.1)
Never or Hardly Ever	62 (2.0) 205 (1.3)	23 (1.9) 258 (1.6)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

*** Sample size insufficient to permit a reliable estimate (see Appendix A).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

A positive relationship between reading textbooks and NAEP U.S. history scores is evident at both grades 4 and 8. At grade 4, students who reported that they read a textbook at least once a week had a higher average score than those who reported doing so less frequently. Students who reported reading a textbook once or twice a month had an average score that was not significantly different from the average scores of students who reported reading a textbook a few times a year, but was significantly higher than the score of those who reported never reading a textbook as part of social studies or history study. At grade 8, students who reported reading a textbook at least once a week had a higher average score than those who reported doing so infrequently (a few times a year) or never. As with grade 4, eighth graders who reported reading a textbook once or twice a month had an average score that was not significantly different from that of students who reported reading a textbook a few times a year, but that was significantly higher than the average score of those who reported never reading a textbook as part of social studies or history study.

The relationship between average U.S. history scores and reading extra written materials, such as biographies and historical fiction, as part of social studies or history classes is the opposite of that observed for reading textbooks. At grade 4, students who reported reading extra written materials monthly outperformed those who reported doing so more frequently. However, students who reported reading extra written materials monthly or a few times a year had higher average scores than those who reported reading such materials about every day or never. Eighth graders who reported reading extra written materials monthly or a few times a year also outperformed those who reported doing so more frequently and those who reported never reading such materials.

The pattern for the use of writings of historical people is similar. Students who reported using such documents monthly or less frequently performed better than students who reported using them more often, at both grades 4 and 8. (Also at grades 4 and 8, students who reported weekly use of writings of historical people outperformed those who reported daily use.) But eighth graders who reported using writings of historical people once or twice a month or a few times a year performed better than students who reported never using such documents. (Those who reported using such materials a few times a year outperformed those who reported doing so once or twice a month.)

TABLE 4.6

Students' Reports on Use of Textbooks, Extra Written Materials, and Primary Documents Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Read Textbooks		
Almost Every Day	46 (1.5) 215 (1.5)	51 (1.3) 262 (0.7)
Once or Twice a Week	25 (0.8) 207 (1.2)	33 (1.1) 260 (0.9)
Once or Twice a Month	8 (0.5) 198 (3.1)	8 (0.6) 258 (2.4)
A Few Times a Year	8 (0.6) 191 (3.0)	5 (0.5) 249 (2.9)
Never	13 (0.8) 183 (2.3)	3 (0.3) 242 (3.8)
Read Extra Written Material		
Almost Every Day	12 (0.5) 198 (2.3)	10 (0.5) 254 (1.5)
Once or Twice a Week	29 (0.9) 206 (1.6)	28 (0.8) 256 (1.0)
Once or Twice a Month	19 (0.6) 213 (1.6)	25 (0.6) 265 (0.9)
A Few Times a Year	15 (0.6) 210 (1.6)	19 (0.6) 267 (1.1)
Never	24 (0.8) 200 (1.8)	17 (0.6) 253 (1.4)
Use Writings of Historical People		
Almost Every Day	10 (0.5) 187 (2.0)	6 (0.3) 242 (2.2)
Once or Twice a Week	17 (0.6) 195 (1.9)	13 (0.5) 253 (1.1)
Once or Twice a Month	19 (0.7) 208 (1.8)	21 (0.6) 262 (1.0)
A Few Times a Year	17 (0.8) 211 (1.8)	24 (0.7) 268 (1.0)
Never	38 (0.9) 213 (1.2)	35 (0.8) 259 (0.8)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

It should be noted that there are some discrepancies between student- and teacher-reported frequencies of use of written materials. It is not possible to offer conclusive reasons for these discrepancies or to determine which report most accurately represents activities in the classrooms. The reports represent the impressions of students and teachers.

Using Maps and Globes. Maps and globes are common instructional aides in fourth- and eighth-grade classrooms. Over one-half of fourth graders (53 percent) and 44 percent of eighth graders reported using maps and globes in social studies or history classes at least once a week. Only 8 percent of fourth graders and 10 percent of eighth graders reported never using these materials.

As shown in Table 4.7, the relationships between the frequency of use of maps and globes in social studies or history classes and average NAEP U.S. history scores indicates that some use of these materials is associated with higher scale scores. At grades 4 and 8, students who reported never using maps or globes had the lowest average scores. Fourth and eighth graders who reported weekly use of maps and globes had higher average scores than those who reported infrequent (a few times a year) use. At grade 8, students who reported monthly use of maps and globes had higher average scores than those who reported using these materials only a few times a year. However, the average scores for fourth- and eighth-grade students who reported using these materials daily were significantly lower than those of students who reported weekly or monthly use of maps and globes.

Watching Movies, Videos, and Filmstrips. The frequency of use of movies, videos, and filmstrips to augment the information presented through written material or teachers' lessons differs between the fourth and eighth grades. At grade 4, nearly half (48 percent) of students reported watching a movie, video, or filmstrip at least once a week. At grade 8, however, movies, videos, and filmstrips were used less frequently — only one-third (33 percent) reported watching these materials at least weekly.

As with the use of maps and globes, some use of these materials is associated with higher scale scores (see Table 4.8). At both grades 4 and 8, students who reported watching these materials daily and those who reported never watching them had lower average scores than those who reported watching them weekly, monthly, or a few times a year. (The average scores for students who reported watching these materials daily and those who reported never watching them were not

TABLE 4.7

Students' Reports on Use of Maps and Globes Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Almost Every Day	21 (0.9) 204 (1.5)	14 (0.7) 257 (1.6)
Once or Twice a Week	32 (0.9) 213 (1.7)	30 (0.7) 264 (0.8)
Once or Twice a Month	25 (0.7) 210 (1.4)	30 (0.8) 263 (0.9)
A Few Times a Year	14 (0.6) 205 (2.2)	17 (0.7) 258 (1.2)
Never	8 (0.6) 188 (3.1)	10 (0.5) 250 (1.5)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

TABLE 4.8

Students' Reports on Watching Movies, Videos, and Filmstrips Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Almost Every Day	18 (0.7) 194 (1.6)	9 (0.5) 246 (2.0)
Once or Twice a Week	30 (1.1) 207 (1.7)	24 (1.1) 260 (1.1)
Once or Twice a Month	32 (0.8) 218 (1.5)	41 (0.9) 264 (0.8)
A Few Times a Year	15 (0.7) 208 (2.4)	18 (0.9) 264 (1.1)
Never	6 (0.4) 189 (2.9)	8 (0.6) 249 (1.9)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

significantly different.) At grade 4, monthly viewing of these materials was associated with a higher average score than less frequent (a few times a year) or weekly viewing. At grade 8, students who reported watching these materials weekly had a lower average score than those who reported watching such materials a few times a year or once or twice a month.

Use of Computers. The use of computers in fourth-grade social studies or history classrooms appears to be common and frequent. Table 4.9 presents students' reports of how often they use computers in their social studies or history classes and related average NAEP U.S. history scores. Approximately one-half (54 percent) of fourth graders reported using a computer at least once a week. At grade 8, however, computers were not used as frequently. About one-half (52 percent) of eighth-grade students reported never using computers in their social studies or history classroom.

TABLE 4.9

Students' Reports on Use of Computers in Social Studies or History Classrooms Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Almost Every Day	19 (1.2) 199 (2.3)	8 (0.6) 246 (2.1)
Once or Twice a Week	35 (1.6) 208 (1.5)	11 (0.5) 248 (1.4)
Once or Twice a Month	14 (0.7) 207 (2.0)	13 (0.7) 258 (1.6)
A Few Times a Year	11 (0.6) 213 (2.5)	17 (0.8) 264 (1.2)
Never	22 (1.0) 211 (1.8)	52 (1.4) 264 (0.8)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

The relationship between use of computers in the classroom and scores on the NAEP U.S. history assessment is negative for both grades 4 and 8. At grade 4, the only significant differences associated with computer use were that students who reported using a computer about every day had a lower average score than those who reported using a computer weekly or those who reported infrequent use (a few times a year) or never using a computer in their social studies or history classroom. At grade 8, a negative relationship between frequency of computer use and average NAEP U.S. history scores was observed. As the reported frequency of computer use increased, the average scores for those students decreased. The only exceptions were no significant difference between daily and weekly use, and between infrequent use (a few times a year) and never using a computer in the social studies or history classroom.

Again, the relationship between use of instructional materials such as computers and scores on the NAEP U.S. history assessment cannot be interpreted in a causal fashion. A variety of factors, including student ability level and the nature of the material being taught, influence a teacher's decision to use a particular method.

Instructional Activities

Teachers of social studies and history augment typical classroom lectures or discussions with a wide variety of activities. Many teachers may ask students to communicate their historical knowledge and understanding through short or extended writing assignments. Homework or extended projects may be designed to reinforce classroom lessons. Finally, field trips or outside speakers may be aimed at bringing history to life for students.

As part of the NAEP assessments, teachers and students were asked to report on a wide array of instructional activities that might take place in their social studies or history classrooms. As with the previous section, the discussion of selected activities will focus on grades 4 and 8.

Incorporating Writing into History or Social Studies Instruction. The importance of writing in the history curriculum has been stressed by many contemporary observers.[2] Teachers have long believed that the ability to write clearly about a subject demonstrates genuine understanding of it. In addition, writing has been associated with critical thinking and with self-motivated learning.[3] Teachers and students in grades 4 and 8 reported how often students were asked to write short answers (a paragraph or less) to questions and reports, as part of their social studies or history study. The results are presented in Tables 4.10 and 4.11. Students and their teachers reported fairly frequent writing of short answers, and less frequent writing of reports.

Table 4.10 presents the teacher-reported data for frequency of students' writing of short answers or reports. As with teacher data about use of materials, generally there were no significant relationships between student scores and frequency of writing short answers or reports. (The exception occurs at grade 4, where weekly writing of reports was associated with a lower average scale score.) Further, as was discussed previously, there are also discrepancies between student- and teacher-reported data.

The frequencies of writing activities are of interest. Teachers of 68 percent of fourth-grade students reported having students write short answers to questions at least once or twice a week. The percentage was higher for eighth graders; teachers of 79 percent of eighth graders reported having their students write short answers to questions at least weekly. The teachers of only 7 percent of fourth graders and 4 percent of eighth graders reported never or hardly ever having students write short answers. Perhaps predictably, given the greater time and effort involved, writing of reports occurred less frequently at both grades than writing short answers to questions. Teachers of only 6 percent of fourth graders and 4 percent of eighth graders reported having their students write reports weekly. The teachers of the majority of students (63 percent at fourth grade and 66 percent at eighth grade) reported having students write reports once or twice a month. Thirty-one percent of fourth graders and 30 percent of eighth graders had teachers who reported never or hardly ever having students write reports.

TABLE 4.10

Teachers' Reports on Student Writing of Short Answers or Reports Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Write Short Answers		
Almost Every Day	9 (1.4) 202 (3.3)	17 (2.1) 258 (1.9)
Once or Twice a Week	59 (2.1) 207 (1.5)	62 (2.7) 261 (0.9)
Once or Twice a Month	25 (1.9) 204 (2.2)	17 (1.9) 263 (2.2)
Never or Hardly Ever	7 (1.2) 203 (4.9)	4 (1.3) 260 (7.3)
Write Reports		
Almost Every Day	0 (0.0) ***	0 (0.2) ***
Once or Twice a Week	6 (1.2) 188 (4.2)	4 (0.9) 256 (4.0)
Once or Twice a Month	63 (2.3) 207 (1.4)	66 (2.5) 261 (0.9)
Never or Hardly Ever	31 (2.1) 205 (2.1)	30 (2.6) 259 (1.6)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

The estimates of population percentages reported as zero (and standard errors reported as 0.0) are actually nonzero, but rounded to zero when reporting to the nearest integer (or nearest tenth in the case of the standard errors).

*** Sample size insufficient to permit a reliable estimate (see Appendix A).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

As can be seen in Table 4.11, students' reports of writing activities in their social studies or history classes differ from those of their teachers. For example, 35 percent of fourth graders reported writing short answers about every day compared to only 9 percent based on their teachers' reports. At least 70 percent of students in grades 4 and 8 reported writing short answers to questions at least once a week. Less than 10 percent reported never doing so.

Generally, students' reports of some writing bears a positive relationship to average U.S. history scale scores. At both grades, students who reported writing short answers to questions a few times a year or never scored significantly lower, on average, than did those who reported doing so at least once a month. At grade 4, students who reported writing short answers at least once a week had higher average scale scores than did those who reported doing so once or twice a month. At grade 8, students who reported writing short answers to questions once or twice a week as part of their social studies or history classes outperformed those who reported doing so about every day.

As for written reports, the vast majority of students said they write reports during the course of the school year. Nearly two-thirds of fourth graders (64 percent) and slightly more than one-half of the eighth graders (52 percent) reported writing reports at least once or twice a month. Fourth- and eighth-grade students who reported writing reports once or twice a month or a few times a year outperformed those who reported doing so more frequently (daily or weekly) and those who reported never writing reports. With one exception, students who reported writing reports about every day had lower average history scale scores than all other groups, including those who reported never writing reports; grade 8 students who reported weekly use did not outperform those reporting daily use. Also at grade 8, students who reported writing reports a few times a year outperformed those who reported doing so once or twice a month, and students who reported never writing reports had a higher average score than did those who reported writing reports once or twice a week.

TABLE 4.11

THE NATION'S REPORT CARD 1994 U.S. History Assessment

Students' Reports on Writing of Short Answers or Reports Grades 4 and 8

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Write Short Answers		
Almost Every Day	35 (0.8) 209 (1.4)	33 (1.0) 259 (0.9)
Once or Twice a Week	35 (0.8) 213 (1.4)	40 (0.8) 263 (0.8)
Once or Twice a Month	15 (0.5) 202 (1.9)	16 (0.6) 262 (1.4)
A Few Times a Year	8 (0.4) 189 (2.9)	5 (0.4) 250 (2.8)
Never	8 (0.4) 187 (2.8)	6 (0.4) 249 (1.9)
Write Reports		
Almost Every Day	8 (0.5) 180 (2.8)	4 (0.3) 240 (2.9)
Once or Twice a Week	16 (0.8) 193 (1.7)	10 (0.6) 245 (1.7)
Once or Twice a Month	40 (0.9) 215 (1.4)	38 (1.0) 260 (0.9)
A Few Times a Year	26 (1.1) 216 (1.8)	35 (1.2) 269 (0.9)
Never	11 (0.6) 198 (2.2)	14 (0.7) 255 (1.4)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Homework. Teachers frequently reinforce the lessons taught in their classrooms with homework. As part of the NAEP assessment, students at all three grades were asked to indicate the amount of time they usually spend on homework each day. This question was not restricted to social studies or history homework but referred to homework across all subject areas. Table 4.12 presents the results for fourth, eighth, and twelfth graders. Overall, there is a positive relationship between spending some time on homework (i.e., homework assigned for any class, not just social studies or history) and average U.S. history scale scores.

TABLE 4.12

Students' Reports on Time Spent on Homework Each Day Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8	Grade 12
	Percentage and Average Scale Score	Percentage and Average Scale Score	Percentage and Average Scale Score
Don't Usually	13 (1.1)	7 (0.6)	13 (0.6)
Have Any	209 (3.3)	245 (1.8)	272 (1.4)
Usually Don't	3 (0.3)	8 (0.4)	8 (0.4)
Do It	180 (3.4)	244 (1.4)	279 (1.6)
1/2 Hour or Less	39 (1.2)	22 (0.8)	23 (0.6)
	204 (1.4)	257 (1.2)	287 (1.0)
1 Hour	30 (0.8)	36 (0.7)	29 (0.5)
	209 (1.5)	262 (0.7)	287 (1.0)
More than	16 (0.8)	27 (0.9)	26 (0.7)
1 Hour	200 (1.7)	266 (0.9)	295 (1.2)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

At all grades, between 7 and 13 percent of the students indicated that they did not usually have homework assigned. For students who indicated that they usually had homework assigned, those reporting that they spent some time on homework significantly outperformed those who reported they did not usually do their homework. At grade 4, the average score for students who reported spending more than one hour on homework each day was significantly less than that for those who reported spending one hour. At grade 8, average scale scores increased significantly with each incremental increase in the time spent on homework. At grade 12, the average scale score for students who reported spending more than one hour on homework each day was significantly more than that for those who reported spending one hour or less.

Eighth graders were also asked about how much history homework they did each week. The results for students who reported that they were taking a U.S. history class are presented in Table 4.13. Students who reported spending no time on history homework had the lowest average scale scores and students who reported that they spent two hours a week on history homework had the highest average scale scores. Within this range, each increase in time spent on history homework was related to an increase in average score. For the 12 percent of the students who reported spending more than two hours a week on history homework, the average score was higher than that of students who spent one-half hour or less, but significantly lower than that of students who spent two hours.

TABLE 4.13

Students' Reports on Time Spent on History Homework Each Week Grade 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Percentage	Average Scale Score
For Grade 8 Students Currently Taking a U.S. History Course		
None	14 (0.9)	253 (1.5)
1/2 Hour	34 (0.9)	258 (1.1)
1 Hour	26 (0.8)	263 (0.9)
2 Hours	15 (0.6)	269 (1.4)
More than 2 Hours	12 (0.8)	264 (1.4)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Field Trips and Libraries. Many teachers can supplement their social studies or history instruction by including field trips and outside speakers. Having students work in the library is another way of augmenting the resources available in the classroom. Table 4.14 presents fourth- and eighth-grade students' reports of how often they go on field trips or have outside speakers, and how often they work in the library. As with most of the other variables discussed thus far, moderate frequencies of these activities are positively related to student performance.

The use of field trips or outside speakers, though not frequent at either grade, is more common at grade 4 than at grade 8. At grade 4, only 14 percent reported never being exposed to those activities in the context of a social studies or history class. At grade 8, almost half the students reported never going on field trips or having outside speakers (45 percent).

At both grades 4 and 8, students reporting field trips or outside speakers at least weekly had lower average scores than those students reporting such activities occurring once or twice a month or a few times a year. However, for both grades students reporting these activities occurring a few times a year performed better than students who reported going on field trips or having outside speakers once or twice a month or never. Additionally, at grade 8 students who reported these activities once or twice a month outperformed those reporting these activities never occurred.

As for working in the library as part of their social studies or history classes, the majority of students at grades 4 and 8 reported working in the library at least a few times a year. At grade 4, 41 percent of the students reported that they work in the library as part of their social studies or history classes at least once a week. In comparison, 15 percent of eighth graders reported working in the library at least once a week.

Fourth-grade students who reported working in the library once or twice a month outperformed both those students who reported working in the library every day and those students who reported never working in the library. Students who reported working in the library about every day had a lower average scale score than those who reported doing so a few times a year or never. Eighth-grade students who reported working in the library once or twice a month, a few times a year or never outperformed those students who reported working in the library at least weekly, and grade 8 students who reported working in the library a few times a year outperformed those students who reported working in the library once or twice a month or never.

TABLE 4.14

Students' Reports on Field Trips or Outside Speakers, and Libraries Grades 4 and 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8
	Percentage and Average Scale Score	Percentage and Average Scale Score
Go On Field Trips or Have Outside Speakers		
Almost Every Day	4 (0.3) 164 (2.9)	2 (0.2) 228 (3.1)
Once or Twice a Week	6 (0.4) 169 (3.3)	3 (0.3) 231 (2.6)
Once or Twice a Month	24 (0.9) 208 (1.5)	11 (0.6) 249 (1.5)
A Few Times a Year	51 (1.1) 215 (1.2)	38 (1.0) 266 (1.0)
Never	14 (0.7) 203 (1.7)	45 (1.4) 262 (0.8)
Work in the Library		
Almost Every Day	7 (0.4) 184 (2.5)	4 (0.3) 242 (2.8)
Once or Twice a Week	34 (1.1) 209 (1.5)	11 (0.7) 249 (1.7)
Once or Twice a Month	19 (0.7) 215 (1.7)	28 (1.0) 261 (1.0)
A Few Times a Year	14 (0.6) 210 (2.4)	28 (1.1) 267 (1.0)
Never	26 (1.0) 203 (1.6)	30 (1.4) 259 (0.8)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Students' Home Support

Home support for learning is as important in social studies and history as it is for other subjects surveyed in NAEP assessments. Variables related to the students' home environment frequently show a strong relationship to performance.[4] The following section examines some of these variables and their relationship to NAEP U.S. history scores.

Discussing Schoolwork at Home. One indication that schoolwork is a priority for students and their families is the extent to which it is discussed at home. When students discuss their schoolwork at home, they establish an important link between home and school. Several recent studies have documented the increased achievement of students whose parents are involved in their schooling.[5] Fostering this valuable link has become a major objective of many recent education reform efforts, including *Goals 2000*,[6] which seeks to increase the cooperation between parents and schools.

Students in the NAEP 1994 U.S. history assessment were asked how frequently they discussed their studies at home. (This question was not restricted to discussing social studies or history studies at home but referred to all subject areas.) Their responses are summarized in Table 4.15. There is a positive relationship between frequency of discussion and average scale scores. At grade 4, 18 percent of the students reported never or hardly ever discussing their studies at home. At grades 8 and 12, the lack of such discussion was reported by 22 and 25 percent of the students, respectively.

The majority of students at each grade reported discussing schoolwork at home at least sometimes. At grade 4, over three-fourths of the students (76 percent) reported daily or weekly discussions. At grades 8 and 12, daily or weekly discussions were reported by more than 60 percent of the students.

TABLE 4.15

Students' Reports on the Frequency with Which They Discuss Their Studies at Home Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8	Grade 12
	Percentage and Average Scale Score	Percentage and Average Scale Score	Percentage and Average Scale Score
Almost Every Day	54 (1.0) 208 (1.1)	38 (0.7) 266 (0.7)	31 (0.7) 292 (1.0)
Once or Twice a Week	22 (0.7) 207 (1.4)	29 (0.6) 262 (1.1)	30 (0.5) 290 (0.9)
Once or Twice a Month	6 (0.3) 206 (2.7)	11 (0.4) 259 (1.2)	14 (0.5) 285 (1.4)
Never or Hardly Ever	18 (0.5) 192 (1.7)	22 (0.6) 252 (1.2)	25 (0.5) 278 (1.0)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

At all grades, students who reported they never or hardly ever discussed schoolwork at home had lower average scores than did those who reported having some regular discussions. Fourth graders' average U.S. history scores did not vary with the frequency of these regular discussions. However, eighth graders who reported daily discussions outperformed those who reported less frequent discussions, and twelfth graders who reported daily or weekly discussions outperformed those who reported monthly discussions.

Access to Literacy Material. Students can also be exposed to historical information through a variety of written sources outside of school. Moreover, parents who read regularly model a pattern of curiosity and lifelong learning that is important for their children's academic success.[7] As part of the NAEP assessments, students are asked about the presence of four different types of literacy material in their homes — magazines, newspapers, encyclopedias, and at least 25 books. The percentages of students reporting that their families have all four types, only three types, or two or fewer types of literacy material are presented in Table 4.16. In general, students who reported having more types of literacy material in their homes also had higher average U.S. history scores.

TABLE 4.16

Students' Reports on the Number of Different Types of Literacy Material in Their Home Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8	Grade 12
	Percentage and Average Scale Score	Percentage and Average Scale Score	Percentage and Average Scale Score
4 Items	37 (1.0) 219 (1.5)	50 (0.9) 267 (0.8)	56 (0.7) 293 (0.7)
3 Items	33 (0.7) 206 (1.3)	30 (0.7) 258 (0.9)	27 (0.5) 285 (1.2)
2 or Fewer Items	30 (0.8) 187 (1.4)	21 (0.7) 243 (1.0)	17 (0.5) 269 (1.2)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Television Viewing Habits. Past NAEP assessments have highlighted the national concern over the amount of time students spend watching television. For example, the 1992 and 1994 NAEP reading assessments found a negative relationship between the amount of television viewing and reading performance. One major concern has been that time spent watching television may be displacing time that could be spent on schoolwork or reading activities.

Students' reports of their television viewing habits are presented in Table 4.17. Clearly, a large amount of students' time is devoted to watching television. The percentages of students who reported watching four or more hours of television each day were 42 percent at grades 4 and 8, and 25 percent at grade 12.

At all grades, more frequent television viewing was associated with lower U.S. history scores. Students who reported watching television three hours or less per day outperformed, on average, those who reported watching four or more hours daily. In addition, at grade 12, students who reported watching television one hour or less per day had a higher average score than did those reporting two to three hours of daily viewing. Across the three grades, students who reported watching four to five hours daily had higher average scale scores than did those who reported watching television six or more hours a day.

TABLE 4.17

Students' Reports on the Amount of Time Spent Watching Television Each Day Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Grade 4	Grade 8	Grade 12
	Percentage and Average Scale Score	Percentage and Average Scale Score	Percentage and Average Scale Score
1 Hour or Less	20 (0.7) 210 (1.8)	13 (0.5) 266 (1.3)	27 (0.7) 293 (0.9)
2 to 3 Hours	38 (0.7) 212 (1.2)	45 (0.8) 264 (0.7)	47 (0.7) 288 (0.9)
4 to 5 Hours	21 (0.7) 205 (1.4)	27 (0.7) 257 (0.9)	18 (0.5) 280 (1.0)
6 Hours or More	21 (0.8) 185 (1.7)	15 (0.5) 245 (1.3)	7 (0.3) 267 (1.8)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages and average scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Summary

The picture of social studies and history education and student performance gained from an examination of background variables relating both to course work and classroom practices is quite varied. Encouragingly, exposure to U.S. history was associated with higher average scale scores for students in grades 4 and 8. While grade 12 students who indicated that they were currently taking U.S. history had lower scores, the more semesters of potential history-related coursework done, the higher the scores.

Aside from textbooks, no single instructional material appeared to have widespread frequent use at grades 4 and 8. Despite the call from history educators for the use of primary historical sources in the classroom, few students and teachers reported using primary source materials. Generally, some (as opposed to frequent) use of instructional materials, such as maps and globes, was associated with higher scale scores, although computer use was negatively related to scores. Similarly, results related to instructional practices, such as library use, for the most part suggested a positive relationship between moderate frequencies and average scale scores.

For those eighth graders who reported that they were currently taking U.S. history, average scale scores increased incrementally in relation to more time spent on history homework (up to more than two hours, at which point scores decreased).

Performance associated with home support variables is consistent with educational research findings. Discussion of schoolwork at home and literacy materials in the home were both associated with higher scores, while frequent television-watching was associated with lower scores. (Given the number of factors the NAEP assessment cannot control for, the reader is again cautioned against drawing causal inferences.)

Endnotes

1. Mahood, Wayne. (1995). Bringing the American Civil War to the classroom: Suggestions and sources. *Social Education, 59*(1), 17-22.

   Nelson, Jennifer T. (1994). Voices from the past: Quotations. *Southern Social Studies Journal, 19*(2), 45-54.

   Van Middendorp, Judy E. and Lee, Sharon. (1994) Literature for children and young adults in a history classroom. *Social Studies, 85*(3), 117-20.

   Wilson, Don W. (1993). Teaching for the future by reaching into the past. *Teaching History: A Journal of Methods, 18*(1), 3-13.

2. Gilstrap, Robert L. (1991). Can writing be used to improve learning and thinking? *Social Studies Review, 31*(1), 26-32.

   Moore, Ellen E. (1990). Writing to learn in the social studies classroom. *New Mexico English Journal 5*(2), 15-18.

   Wills, Howard. (1993). *Writing is learning: Strategies for math, science, social studies, and language arts.* Bloomington, IN: EDINFO Press.

3. Capps, Kline, and Vocke, David E. (1991). Developing higher-level thinking skills through American history writing assignments. *OAH Magazine of History 6*(2), 6-9.

   Risinger, Frederick C. (1987). *Improving writing skills through social studies.* ERIC Clearinghouse for Social Studies/Social Science Education, ERIC Digest 40. Washington, DC: Office of Educational Research and Improvement.

4. Christenson, S.L. (1992). Family factors and student achievement: An avenue to increase students' success. *School Psychology Quarterly, 7*(3). 178-206.

5. Christenson, S.L. (1992). Ibid.

   Heller, L.R., and Fantuzzo, J.W. (1993). Reciprocal peer tutoring and parent partnership: Does parent involvement make a difference? *School Psychology Review, 22*(3), 517-34.

6. *The National Educational Goals Report: Building a Nation of Learners.*(1994). (Report of the National Education Goals Panels: U.S. Government Printing Office).

7. Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* (New York, NY: Oxford University Press).

CHAPTER 5

What Students Know and Can Do in U.S. History

This chapter provides additional perspectives on what students know and can do in U.S. history, by examining the specific knowledge and skills demonstrated by students at various points on the composite history scale, and the four subscales (the subscales correspond to the four U.S. history themes, around which the assessment was organized).

An overview of the U.S. historical knowledge and abilities demonstrated by students within three performance ranges on the U.S. history scale, and information on their self-reported study habits, is included. Average scale score results are presented for the nation and for selected subgroups of students based on the four subscales. Finally, selected illustrative questions are displayed in item maps, which indicate the types of questions that were likely to be answered successfully by students scoring at particular levels on the thematic subscales.

Overview of Students' Performance on the NAEP U.S. History Scale

The NAEP U.S. history scale, ranging from 0 to 500, summarizes the overall scores of students at the fourth, eighth, and twelfth grades. The following descriptions of students' abilities are based on sets of questions that were answered successfully by students performing in three ranges on the scale. These ranges represent lower, middle, and higher performance based on percentile distribution. The sets of questions identified in each of the three ranges on the scale were analyzed by history education experts to characterize the nature of students' history knowledge and abilities. The procedures used to generate this portrait of students are described in Appendix B.

Fourth Grade Profile. Fourth graders who were near the 25th percentile (scale range 171 to 187) demonstrated that they could recall major figures and events that are very well known in American culture, such as those connected with national holidays or landmark events. They recognized direct clues from pictorial sources, and most of their recall was tied to such sources.

Fourth graders who were near the 50th percentile (scale range 205 to 215) were able to recall some major historical figures and events as well as some concepts, such as the goals of major reform movements. These students also used several types of sources, such as drawings, time lines, and texts, to make simple historical inferences, such as connecting clues in a photograph to knowledge of its context. They demonstrated a developing ability to grasp historical concepts such as the relationships between geography and human settlement, or comparisons across time.

Fourth graders who were near the 90th percentile (scale range 246 to 263) were able to recall a range of major historical figures, events, and concepts. These students also had some ability to use their knowledge to make historical connections, such as comparisons across time or connections between prior knowledge and new information. They could comprehend information from a variety of primary and secondary sources, including time lines, diaries, pictorial sources, and texts. In addition, they displayed rudimentary map-reading skills; for example, they could use maps together with prior knowledge to answer a historical question. These students had a developing sense of how primary sources might be used and a developing awareness of cause-and-effect relationships.

Eighth Grade Profile. Eighth graders who were near the 25th percentile (scale range 233 to 244) could recall some major historical figures and events, such as a site of U.S. military involvement or the cause associated with a famous individual. These students had some grasp of the appropriate use of primary and secondary historical sources. In addition, they could use clear clues found in pictorial documents, such as photographs or engravings, to draw direct inferences.

Eighth graders who were near the 50th percentile (scale range 257 to 265) could recall historical figures, events, and concepts, and could identify some details of the context of what they recalled, such as the general chronology of important events. These students could combine knowledge with clues from both visual and written historical sources to draw direct historical inferences. They could identify or recognize cause-and-effect relationships and demonstrated some ability to understand and interpret charts and graphs presenting numerical information. They also demonstrated a limited ability to recognize a point of view in primary source material.

Eighth graders who were near the 90th percentile (scale range 292 to 308) could recall a fairly broad range of political, social, and economic events, topics, and themes. The knowledge demonstrated went beyond the surface level in many cases. These students demonstrated some understanding of historical context and chronology and could apply their knowledge to make connections to historical generalizations, patterns, or themes. Given data, these students could draw cause-and-effect and comparative relationships. They could understand and interpret a variety of maps, charts, graphs, and draw conclusions from some of them. They were able to use a variety of primary sources to answer historical questions and could generally recognize the point of view evident in these sources.

Twelfth Grade Profile. Twelfth graders who were near the 25th percentile (scale range 259 to 270) could recall major historical figures and events and were familiar with a few landmark documents. They were able to draw straightforward inferences from a variety of documents, particularly texts and pictorial sources, and could read and understand straightforward data tables. These students had some grasp of chronology and historical concepts such as cause and effect; in some cases these students could make historical connections for themselves.

Middle-performing twelfth graders who were near the 50th percentile (scale range 284 to 292) demonstrated knowledge of general historical chronology, particularly for the twentieth century. They were able to extract information from a wide range of documents, including political cartoons, maps, texts, and pictorial sources. They showed some ability to analyze the point of view evident in various kinds of documents. These students were also able in some cases to bring a base of historical knowledge to bear on interpreting a source. They demonstrated a beginning understanding of historical concepts, such as the ability to compare and contrast situations across time and to identify cause-and-effect relationships.

Twelfth graders who were near the 90th percentile (scale range 319 to 335) showed a working knowledge of a broad range of events, movements, policies, and concepts in U.S. history. Their knowledge covered political, social, economic, and foreign policy questions, and was strongest for the nineteenth and twentieth centuries. Their understanding frequently went beyond the surface level. For example, these students demonstrated some understanding of the impact of major religious and social movements. They were able to connect new material provided in the question to prior knowledge, to make inferences and generalizations, and to draw conclusions.

Higher-performing twelfth graders also could work with a wide variety of historical sources and were able to evaluate points of view. They were able to comprehend and interpret fairly complex primary sources, including some written in formal, archaic language. These students showed an ability to evaluate connections between events and policies, to use historical reasoning in considering a historical issue, and to grasp historical significance.

Profiles of Students' U.S. History Knowledge, Abilities, and Study Habits

The following three figures (5.1, 5.2, and 5.3) represent profiles of the lower-, middle-, and higher-performing students. The profiles link the knowledge and abilities of these students with their self-reported study habits. The knowledge and abilities presented in these figures summarize the performance descriptions from the previous section. The study habits presented in the figures are based on students' self-reports concerning three activities: number of pages read per day in school, hours per day spent on homework, and frequency of discussing studies at home.

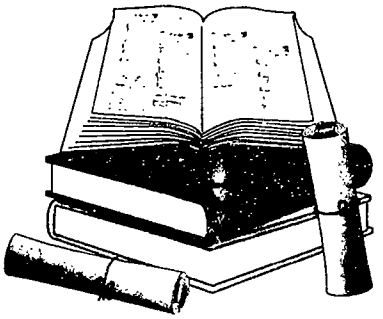
By examining all three profiles, a common pattern emerges at each grade with one exception: at the fourth-grade 90th percentile. Students performing at the higher range of the scale were more likely to read more pages per day in school, spend more time on homework, and more frequently discuss their studies at home.[1]

Figure 5.1 Profiles of Lower-, Middle-, and Higher-Performing Fourth Graders: U.S. History Knowledge, Abilities, and Study Habits

25th Percentile

Fourth-grade students who were near the 25th percentile could:

- recall famous historical figures
- recall events connected to national holidays
- recognize pictorial sources



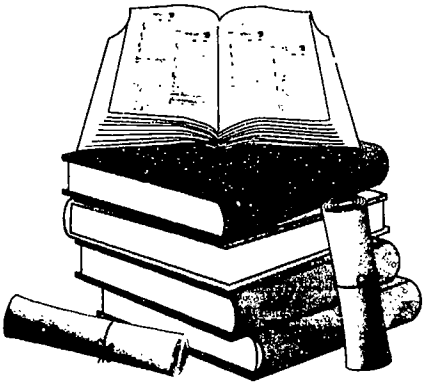
Study habits of fourth graders who were near the 25th percentile:

- 33 percent read more than 15 pages each day in school and for homework
- 43 percent spent one or more hours each day on homework
- 71 percent discussed studies at home at least once or twice a week

50th Percentile

Fourth-grade students who were near the 50th percentile could:

- recall some major concepts
- use various sources, such as time lines and texts, to make simple historical inferences
- begin to grasp historical concepts, such as relationships between geography and human settlement



Study habits of fourth graders who were near the 50th percentile:

- 38 percent read more than 15 pages each day in school and for homework
- 47 percent spent one or more hours each day on homework
- 76 percent discussed studies at home at least once or twice a week

90th Percentile

Fourth-grade students who were near the 90th percentile could:

- connect prior knowledge and new information
- interpret information from primary sources such as diaries and maps
- begin to grasp cause-and-effect relationships

Study habits of fourth graders who were near the 90th percentile:

- 44 percent read more than 15 pages each day in school and for homework
- 46 percent spent one or more hours each day on homework
- 82 percent discussed studies at home at least once or twice a week

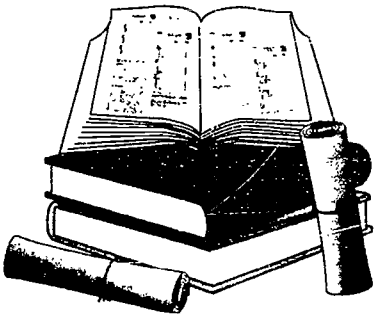
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Figure 5.2 Profiles of Lower-, Middle-, and Higher-Performing Eighth Graders: U.S. History Knowledge, Abilities, and Study Habits

25th Percentile

Eighth-grade students who were near the 25th percentile could:

- recall major historical figures and events
- make basic use of primary and secondary historical sources
- make direct inferences from pictorial sources



Study habits of eighth graders who were near the 25th percentile:

- 16 percent read more than 15 pages each day in school and for homework
- 58 percent spent one or more hours each day on homework
- 61 percent discussed studies at home at least once or twice a week

50th Percentile

Eighth-grade students who were near the 50th percentile could:

- place people, events, and concepts in general historical context
- make direct historical inferences from text and visual sources
- recognize cause-and-effect relationships

Study habits of eighth graders who were near the 50th percentile:

- 21 percent read more than 15 pages each day in school and for homework
- 65 percent spent one or more hours each day on homework
- 66 percent discussed studies at home at least once or twice a week

90th Percentile

Eighth-grade students who were near the 90th percentile could:

- recall a broad range of topics and themes
- apply knowledge of historical context and chronology to make connections to historical patterns and themes
- use data to infer cause-and-effect and comparative relationships

Study habits of eighth graders who were near the 90th percentile:

- 25 percent read more than 15 pages each day in school and for homework
- 71 percent spent one or more hours each day on homework
- 76 percent discussed studies at home at least once or twice a week

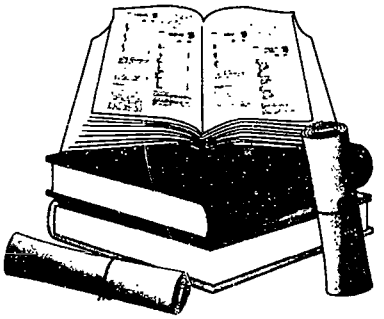
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Figure 5.3 Profiles of Lower-, Middle-, and Higher-Performing Twelfth Graders: U.S. History Knowledge, Abilities, and Study Habits

25th Percentile

Twelfth-grade students who were near the 25th percentile could:

- recognize some major historical documents
- make direct inferences from a variety of documents
- draw simple cause-and-effect relationships



Study habits of twelfth graders who were near the 25th percentile:

- 18 percent read more than 15 pages each day in school and for homework
- 49 percent spent one or more hours each day on homework
- 55 percent discussed studies at home at least once or twice a week

50th Percentile

Twelfth-grade students who were near the 50th percentile could:

- show general knowledge of historical chronology, especially 20th century history
- interpret a wide range of historical documents, including cartoons, maps, and texts
- show some grasp of point of view in documents
- apply prior knowledge to source interpretation

Study habits of twelfth graders who were near the 50th percentile:

- 25 percent read more than 15 pages each day in school and for homework
- 56 percent spent one or more hours each day on homework
- 62 percent discussed studies at home at least once or twice a week

90th Percentile

Twelfth-grade students who were near the 90th percentile could:

- show knowledge of a wide range of movements, policies, and concepts, including 19th and 20th century economic and foreign policy issues
- grasp the impact of major social and religious movements
- interpret and evaluate points of view in various historical sources
- evaluate connections between events and policies

Study habits of twelfth graders who were near the 90th percentile:

- 41 percent read more than 15 pages each day in school and for homework
- 64 percent spent one or more hours each day on homework
- 73 percent discussed studies at home at least once or twice a week

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Average Performance by Theme

As described in Chapter 1, the NAEP 1994 U.S. history assessment was organized around four themes, designed to ensure coverage of the major branches of historical study. The themes were envisioned as broad categories that would necessarily overlap at times. The historical skills outlined in the *NAEP 1994 U.S. History Framework* cut across all the themes.

1. Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

This theme concerns the development of American political democracy from colonial times to the present It covers political events that shaped American democracy, such as the American Revolution, the Civil War, and the fight for civil rights, as well as the core ideas and principles that underlie our institutions. This theme ensures that students' knowledge of the founding of the nation, the writing of the Constitution, and other fundamental components of the nation's political history will be assessed. At the same time, it calls for evaluating students' understanding of the role that major political ideas and conflicts have played at different points in our history.

2. The Gathering and Interactions of Peoples, Cultures, and Ideas

The second theme is broadly defined because it covers a vast component of U.S. history: the interactions among the people and cultures of many countries, racial groups, and religious traditions that have contributed to the development of American society. This theme covers the nature and role of immigration throughout our history, cultural developments, patterns of social organization, and changing roles of men and women.

3. Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

This theme focuses on the economic history of the nation and its development from a rural, agricultural society to an urban, industrialized superpower. It covers the role that economic ideas and beliefs have played in this change as well as the roles of geography and of developments in science and technology.

4. The Changing Role of America in the World

This theme calls for coverage of the many factors — physical geography, political ideals, economic interests, public opinion — that have shaped American foreign policy. It also addresses specific interactions between the United States and other nations and the domestic consequences of developments in foreign policy.

Table 5.1 presents the average scores by thematic subscale for grades 4, 8, and 12. The reporting metrics for the four subscales were created independently of one another and are not generally comparable. Therefore, statements involving comparisons across subscales (i.e., fourth graders performed better on Theme 1 than on Theme 2) are not necessarily meaningful. The reader is referred to Appendix A (page 81) for a more extensive discussion of the U.S. history scaling procedures.

As one might expect, the patterns of performance among subgroups on the composite scale were also generally evident for the thematic subscales.

TABLE 5.1

Average Scale Scores at Various Percentiles by Historical Themes Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Composite	Theme 1 *American Democracy*	Theme 2 *Peoples, Cultures, and Ideas*	Theme 3 *Economic and Technological Changes*	Theme 4 *World Role*
Grade 4					
Average Score	205 (1.0)	208 (1.1)	203 (1.3)	204 (1.3)	204 (1.1)
10th Percentile	14? (2.4)	156 (2.7)	137 (2.1)	139 (2.0)	156 (1.8)
25th Percentile	180 (1.3)	185 (1.6)	174 (1.4)	177 (1.6)	182 (1.3)
50th Percentile	210 (0.9)	212 (1.4)	208 (1.4)	210 (1.6)	207 (1.6)
75th Percentile	234 (1.2)	235 (0.8)	236 (1.5)	238 (1.4)	229 (1.4)
90th Percentile	253 (1.2)	253 (1.1)	258 (1.8)	259 (2.2)	247 (1.1)
Grade 8					
Average Score	259 (0.6)	256 (0.7)	263 (0.7)	259 (0.7)	258 (0.9)
10th Percentile	217 (1.0)	212 (1.3)	221 (1.1)	217 (1.2)	209 (1.6)
25th Percentile	239 (0.9)	234 (0.9)	243 (0.8)	239 (1.0)	234 (1.2)
50th Percentile	261 (0.8)	258 (0.7)	265 (0.9)	261 (0.9)	260 (0.9)
75th Percentile	282 (0.8)	280 (0.9)	286 (0.7)	280 (0.7)	284 (1.0)
90th Percentile	299 (0.8)	299 (1.0)	303 (0.8)	297 (1.1)	305 (1.3)
Grade 12					
Average Score	286 (0.8)	286 (1.1)	284 (0.7)	287 (0.7)	287 (1.0)
10th Percentile	243 (1.0)	229 (1.4)	249 (1.0)	248 (1.1)	240 (1.6)
25th Percentile	265 (0.9)	258 (1.7)	267 (1.0)	268 (0.7)	263 (1.4)
50th Percentile	288 (0.8)	288 (1.3)	286 (0.8)	289 (0.7)	289 (1.0)
75th Percentile	309 (1.0)	316 (1.4)	303 (0.7)	307 (0.8)	313 (1.0)
90th Percentile	326 (0.8)	339 (1.5)	316 (0.9)	322 (0.9)	333 (1.0)

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Performance on Thematic Subscales by Race/Ethnicity. Table 5.2 displays the subscale results as well as the overall average scale score results for racial/ethnic groups. (Racial/ethnic classifications are based on self-reported information provided by students.)

For the most part, the significant differences among racial/ethnic groups on the composite U.S. history scale were also significant on each of the subscales. For example, White and Asian students outperformed their Black and Hispanic peers at all three grades.

However, not all the significant differences among racial/ethnic subgroups observed with the composite NAEP U.S. history scores were observed for each of the four subscales. At all three grades, different results were observed across some racial/ethnic subgroups when average scale scores were examined for the historical themes.

- At grade 4, Hispanic students outperformed Black students on the fourth theme, World Role. On the other three thematic subscales and on the composite U.S. history scale, the performance of the two groups was not significantly different. Also,

TABLE 5.2

Average Scale Scores in Themes by Race/Ethnicity Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Composite	Theme 1 *American Democracy*	Theme 2 *Peoples, Cultures, and Ideas*	Theme 3 *Economic and Technological Changes*	Theme 4 *World Role*
Grade 4					
Total	205 (1.0)	208 (1.1)	203 (1.3)	204 (1.3)	204 (1.1)
Race/Ethnicity					
White	215 (1.2)	216 (1.2)	214 (1.5)	217 (1.5)	213 (1.2)
Black	177 (1.6)	190 (1.9)	171 (2.0)	172 (2.1)	177 (1.9)
Hispanic	180 (2.7)	184 (2.8)	177 (3.2)	175 (3.2)	188 (2.5)
Asian	209 (4.6)	213 (4.1)	206 (6.2)	206 (6.2)	211 (3.4)
Pacific Islander	200 (5.9)	206 (5.6)	196 (7.5)	197 (7.4)	204 (6.5)
American Indian	190 (6.1)	190 (7.5)	188 (6.9)	190 (6.0)	195 (4.9)
Grade 8					
Total	259 (0.6)	256 (0.7)	263 (0.7)	259 (0.7)	258 (0.9)
Race/Ethnicity					
White	267 (0.8)	263 (0.8)	270 (0.9)	266 (1.0)	267 (1.0)
Black	239 (1.4)	238 (1.6)	244 (1.6)	238 (1.6)	231 (2.2)
Hispanic	243 (1.3)	240 (1.6)	249 (1.5)	239 (1.6)	241 (1.6)
Asian	270 (3.6)	268 (3.8)	271 (3.4)	268 (3.6)	271 (4.7)
Pacific Islander	252 (7.1) !	249 (8.6) !	256 (6.7) !	247 (6.9) !	252 (7.8) !
American Indian	246 (3.7) !	242 (3.9) !	250 (3.9) !	245 (3.3) !	246 (4.8) !
Grade 12					
Total	286 (0.8)	286 (1.1)	284 (0.7)	287 (0.7)	287 (1.0)
Race/Ethnicity					
White	292 (0.8)	293 (1.3)	289 (0.7)	293 (0.7)	294 (1.0)
Black	265 (1.5)	262 (2.4)	269 (1.3)	267 (1.4)	262 (2.0)
Hispanic	267 (1.6)	258 (1.9)	269 (1.5)	270 (1.4)	269 (2.3)
Asian	287 (4.0)	287 (5.7)	284 (3.5)	286 (3.6)	292 (4.6)
Pacific Islander	280 (3.9)	272 (5.5)	278 (3.6)	283 (3.9)	286 (4.5)
American Indian	279 (4.0) !	276 (5.6) !	278 (3.8) !	281 (3.7) !	283 (5.2) !

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

! Interpret with caution any comparisons involving this statistic. The nature of the sample does not allow for accurate determination of the variability of this value.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

the differences in performance between Asian and Hispanic students on the composite and other subscales were not in evidence for the World Role theme. Finally, American Indian students outperformed Black students on the World Role theme. The difference between Pacific Islander and Hispanic students noted on the composite scale was significant only for the American Democracy subscale, where Pacific Islander students outperformed Hispanic students. However, Pacific Islanders outperformed Blacks on all scales except the American Democracy subscale.

- At grade 8, Hispanic students outperformed Black students on the World Role theme.
- At grade 12, the differences between Pacific Islander and Black and Hispanic students noted on the composite scale were significant only for the Economic and Technological Changes and World Role subscales. In addition, White students outperformed Pacific Islander students only on the American Democracy subscale.

Performance on Thematic Subscales by Gender. Male and female students' performance on the four thematic subscales is presented in Table 5.3.

The performance differences between males and females varied by subscale. On the composite scale, the only significant difference observed was at grade 12, where males outperformed females. When performance is examined at the subscale level, however, more differences emerge.

- At grades 4, 8, and 12, males outperformed females on the World Role theme.
- At grades 8 and 12, females outperformed males on the People, Cultures, and Ideas theme. At grade 8, females outperformed males on the Economic and Technological Changes theme.
- At all three grades, no significant differences were observed between males and females on the American Democracy theme.

TABLE 5.3

Average Scale Scores in Themes by Gender Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Composite	Theme 1 *American Democracy*	Theme 2 *Peoples, Cultures, and Ideas*	Theme 3 *Economic and Technological Changes*	Theme 4 *World Role*
Grade 4					
Total	205 (1.0)	208 (1.1)	203 (1.3)	204 (1.3)	204 (1.1)
Gender					
Male	203 (1.5)	207 (1.4)	201 (1.9)	202 (1.7)	207 (1.7)
Female	206 (1.1)	209 (1.3)	205 (1.4)	207 (1.5)	202 (1.1)
Grade 8					
Total	259 (0.6)	256 (0.7)	263 (0.7)	259 (0.7)	258 (0.9)
Gender					
Male	259 (0.8)	256 (0.9)	261 (0.9)	256 (1.1)	263 (1.0)
Female	259 (0.7)	256 (0.8)	265 (0.7)	261 (0.8)	253 (0.9)
Grade 12					
Total	286 (0.8)	286 (1.1)	284 (0.7)	287 (0.7)	287 (1.0)
Gender					
Male	288 (0.8)	287 (1.2)	282 (0.8)	287 (0.8)	294 (1.1)
Female	285 (0.9)	284 (1.2)	286 (0.8)	287 (0.9)	281 (1.0)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Performance on Thematic Subscales by Parents' Level of Education. Performance by student-reported parental education level on the four thematic subscales is presented in Table 5.4.

For the most part, the positive relationship between parents' level of education and performance noted with the composite U.S. history scale was also evident with each of the four thematic subscales. As the level of parental education increased, average scale scores increased. The one exception to this pattern that was noted for the composite scale was at grade 4, where there was no significant scale score difference between students who reported that at least one parent had graduated from college and those who reported that at least one parent had some education after high school. The same pattern at grade 4 was observed on each of the subscales.

TABLE 5.4

Average Scale Scores in Themes by Parents' Highest Education Level Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Composite	Theme 1 *American Democracy*	Theme 2 *Peoples, Cultures, and Ideas*	Theme 3 *Economic and Technological Changes*	Theme 4 *World Role*
Grade 4					
Total	205 (1.0)	208 (1.1)	203 (1.3)	204 (1.3)	204 (1.1)
Parents' Education Level					
Graduated College	216 (1.2)	218 (1.2)	214 (1.6)	217 (1.4)	213 (1.2)
Some Education after High School	214 (2.1)	216 (2.3)	212 (2.8)	216 (3.2)	210 (2.0)
Graduated High School	197 (1.8)	201 (2.7)	195 (2.1)	195 (2.8)	200 (2.4)
Did Not Finish High School	177 (3.3)	184 (4.3)	169 (4.0)	174 (3.9)	188 (2.8)
I Don't Know	195 (1.4)	198 (1.3)	193 (1.7)	194 (1.8)	196 (1.6)
Grade 8					
Total	259 (0.6)	256 (0.7)	263 (0.7)	259 (0.7)	258 (0.9)
Parents' Education Level					
Graduated College	270 (0.8)	268 (0.9)	273 (0.8)	269 (1.0)	270 (1.1)
Some Education after High School	264 (0.8)	262 (0.9)	267 (1.0)	263 (1.2)	263 (1.1)
Graduated High School	251 (0.8)	246 (1.0)	255 (1.0)	252 (1.0)	249 (1.4)
Did Not Finish High School	241 (1.3)	239 (1.8)	246 (1.4)	241 (1.9)	237 (2.1)
I Don't Know	238 (1.4)	234 (1.5)	244 (1.4)	237 (1.5)	237 (2.1)
Grade 12					
Total	286 (0.8)	286 (1.1)	284 (0.7)	287 (0.7)	287 (1.0)
Parents' Education Level					
Graduated College	296 (0.9)	299 (1.2)	292 (0.9)	295 (1.0)	298 (1.1)
Some Education after High School	287 (1.2)	287 (1.7)	285 (1.0)	288 (1.0)	287 (1.4)
Graduated High School	276 (1.1)	272 (1.5)	276 (0.9)	279 (0.9)	277 (1.7)
Did Not Finish High School	263 (1.4)	255 (2.7)	266 (1.4)	266 (1.3)	265 (1.7)
I Don't Know	256 (2.7)	246 (4.0)	258 (2.1)	260 (2.9)	259 (3.6)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Performance on Thematic Subscales by Type of Location. Table 5.5 displays the subscale results as well as the overall average scale score results by type of school location.

For the most part, the significant differences among students attending schools in central city, urban fringe/large town, and rural/small town locations on the composite U.S. history scale were also significant on each of the subscales. The exceptions were:

- At grade 4, there was no significant difference between urban fringe/large town students and rural/small town students for the Peoples, Cultures, and Ideas theme.
- At grade 8, there were no significant differences for the Economic and Technological Changes theme. Also, urban fringe/large town students did not outperform rural/small town students on the Peoples, Cultures, and Ideas theme, as they did on the composite scale.
- At grade 12, the difference between rural/small town students and central city students noted for the composite scale was not evident for the Economic and Technological Changes theme.

TABLE 5.5

Average Scale Scores in Themes by Type of Location Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Composite	Theme 1 *American Democracy*	Theme 2 *Peoples, Cultures, and Ideas*	Theme 3 *Economic and Technological Changes*	Theme 4 *World Role*
Grade 4					
Total	205 (1.0)	208 (1.1)	203 (1.3)	204 (1.3)	204 (1.1)
Type of Location					
Central City	198 (2.0)	204 (2.0)	195 (2.4)	196 (2.4)	199 (1.6)
Urban Fringe/Large Town	211 (1.6)	213 (1.6)	209 (2.0)	212 (1.8)	210 (1.5)
Rural/Small Town	203 (2.8)	204 (3.2)	203 (3.8)	202 (3.1)	202 (2.4)
Grade 8					
Total	259 (0.6)	256 (0.7)	263 (0.7)	259 (0.7)	258 (0.9)
Type of Location					
Central City	257 (1.3)	255 (1.5)	262 (1.3)	256 (1.5)	255 (1.6)
Urban Fringe/Large Town	262 (1.2)	260 (1.3)	266 (1.2)	260 (1.4)	263 (1.3)
Rural/Small Town	258 (1.3)	254 (1.5)	261 (1.4)	260 (1.5)	256 (2.2)
Grade 12					
Total	286 (0.8)	286 (1.1)	284 (0.7)	287 (0.7)	287 (1.0)
Type of Location					
Central City	286 (1.3)	287 (1.9)	285 (1.0)	287 (1.2)	287 (1.7)
Urban Fringe/Large Town	289 (1.2)	289 (1.5)	286 (1.1)	289 (1.2)	291 (1.3)
Rural/Small Town	281 (1.0)	280 (1.4)	280 (1.0)	284 (1.0)	282 (1.4)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Performance on Thematic Subscales by Type of School. Table 5.6 displays the subscale results as well as the overall average scale score results by type of school (public and nonpublic). As with the composite scale, students attending nonpublic schools outperformed those attending public schools on all four thematic subscales.

TABLE 5.6

Average Scale Scores in Themes by Type of School Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Composite	Theme 1 *American Democracy*	Theme 2 *Peoples, Cultures, and Ideas*	Theme 3 *Economic and Technological Changes*	Theme 4 *World Role*
Grade 4					
Total	205 (1.0)	208 (1.1)	203 (1.3)	204 (1.3)	204 (1.1)
Type of School					
Public Schools	203 (1.2)	206 (1.3)	201 (1.4)	202 (1.5)	203 (1.2)
Nonpublic Schools	222 (1.9)	226 (2.0)	219 (2.3)	224 (2.3)	219 (1.9)
Catholic Schools	221 (2.5)	224 (2.4)	216 (3.4)	225 (2.8)	217 (2.4)
Other Nonpublic Schools	224 (3.1)	229 (3.3)	224 (4.1)	222 (3.9)	223 (2.6)
Grade 8					
Total	259 (0.6)	256 (0.7)	263 (0.7)	259 (0.7)	258 (0.9)
Type of School					
Public Schools	257 (0.7)	254 (0.8)	261 (0.8)	257 (0.8)	256 (0.9)
Nonpublic Schools	278 (1.1)	276 (1.4)	280 (1.4)	277 (1.5)	278 (1.6)
Catholic Schools	279 (1.5)	277 (1.9)	280 (1.9)	280 (1.8)	278 (2.0)
Other Nonpublic Schools	277 (2.1)	275 (2.4)	280 (2.2)	274 (2.2)	277 (2.7)
Grade 12					
Total	286 (0.8)	286 (1.1)	284 (0.7)	287 (0.7)	287 (1.0)
Type of School					
Public Schools	284 (0.8)	284 (1.2)	283 (0.8)	286 (0.8)	286 (1.1)
Nonpublic Schools	299 (1.3)	303 (1.9)	295 (1.2)	297 (1.2)	300 (1.7)
Catholic Schools	298 (2.2)	301 (3.2)	295 (1.8)	296 (2.1)	301 (2.6)
Other Nonpublic Schools	299 (2.2)	305 (3.2)	295 (1.9)	298 (1.8)	299 (3.1)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated scale scores appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Profiling Students' Performance on the Four Historical Themes

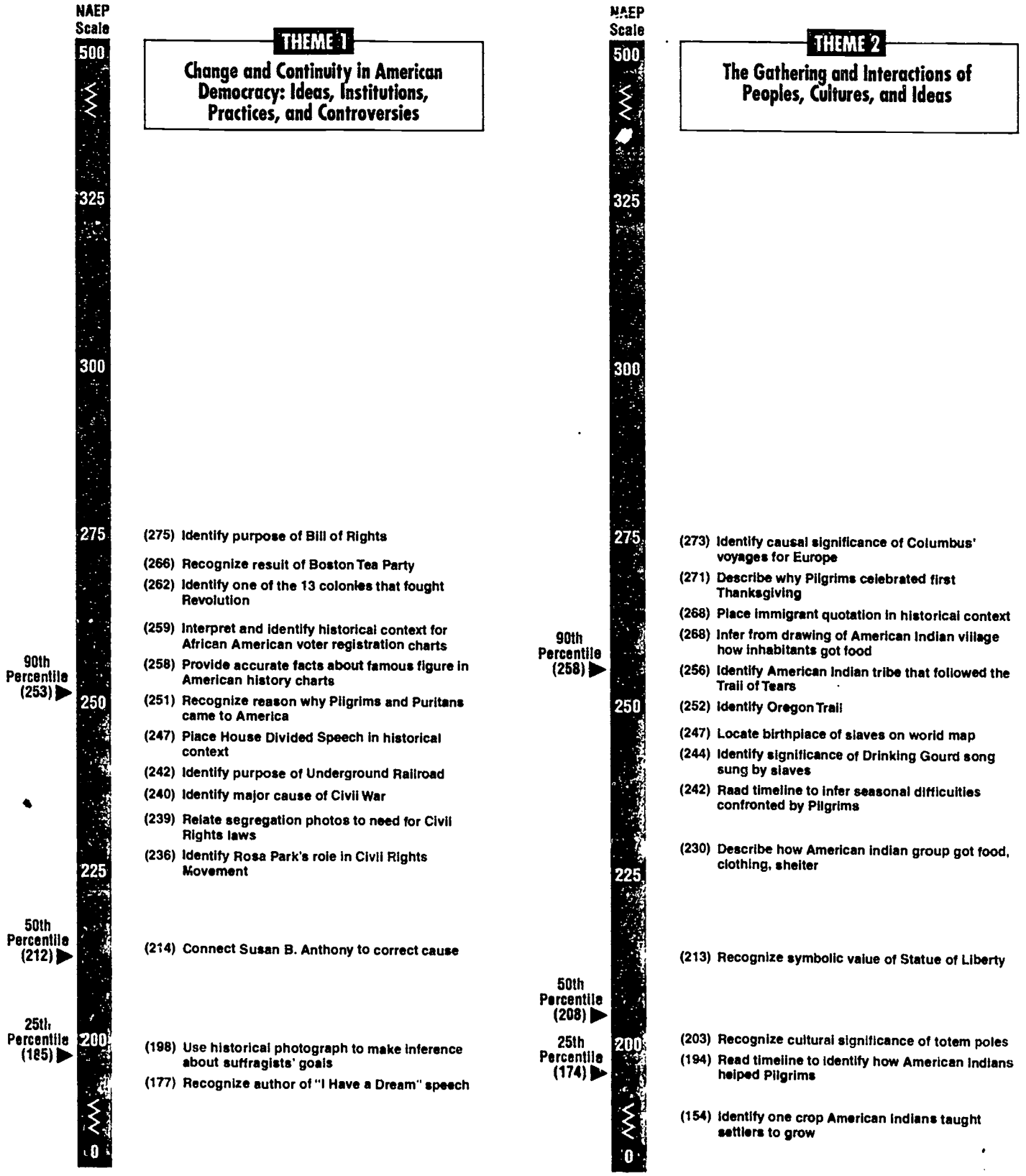
To illustrate the range of assessment tasks that students were asked to perform, this section provides a visual representation of each thematic subscale, called an "item map." The item map shows tasks in the order of their difficulty for the population assessed. Each task is identified by a brief description. Clearly, not all the tasks in the assessment could be presented in these figures. Instead, sample tasks were selected to illustrate the range of abilities across the subscales and to demonstrate the types of tasks that students with different subscale scores could typically answer correctly.

The item maps identify where, on each thematic subscale, individual tasks are answered correctly by approximately two-thirds (65 percent) of the students. (The criterion was set at 74 percent for multiple-choice questions to adjust for the possibility of students answering correctly by guessing.) The point on the subscale item map at which a task is positioned represents the subscale score attained by students who had a 65-percent probability of successfully performing the task. Thus it can be said for each task and its corresponding subscale score that at least 65 percent of students at or above that point on the subscale have adequately performed that task.

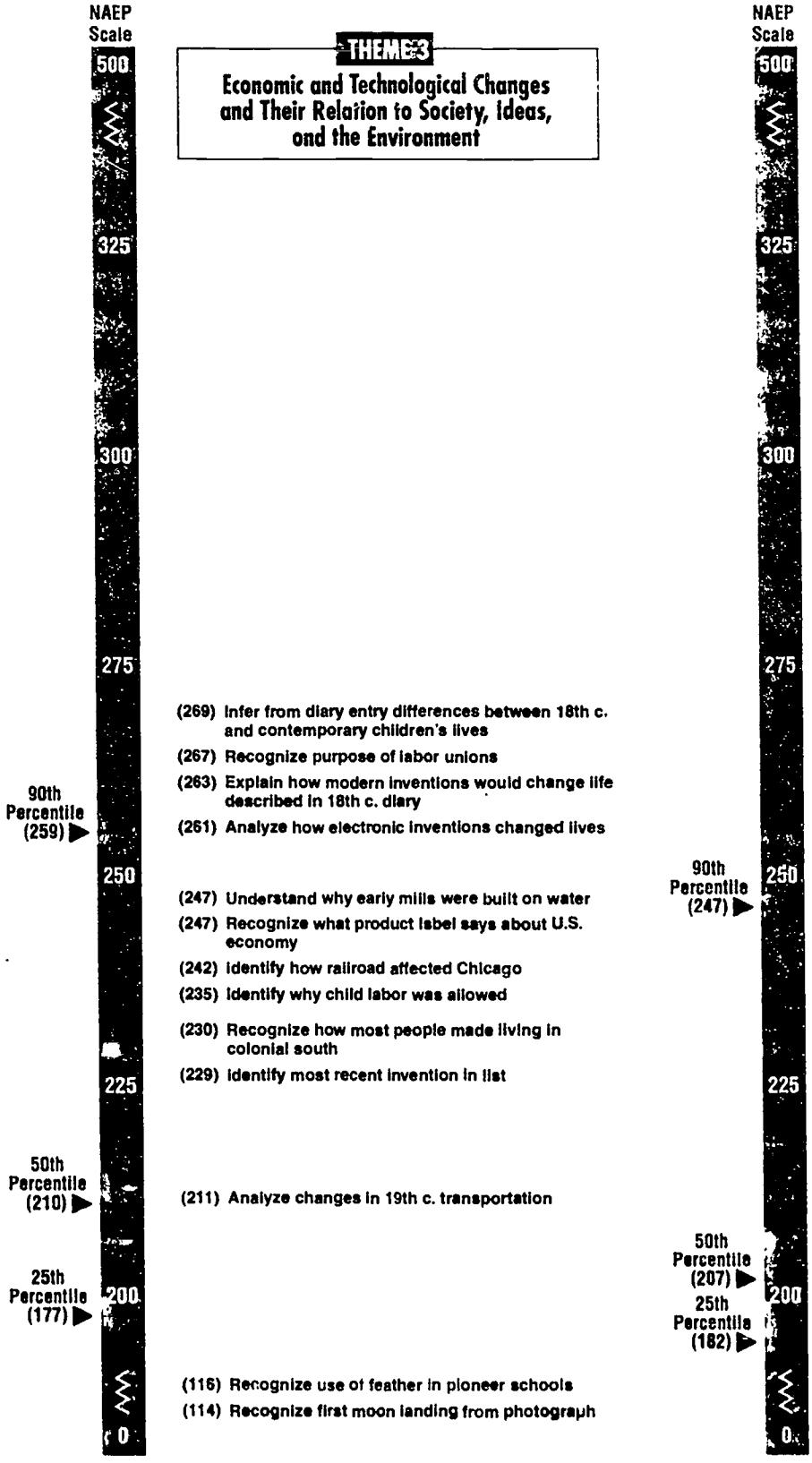
For example, looking at the fourth-grade American Democracy Item Map (Figure 5.4), at least 65 percent of fourth graders with a score of 236 or better on the American Democracy subscale were able to identify Rosa Park's role in the Civil Rights movement. Fourth graders who scored higher than 236 on this subscale were even more likely to be able to answer this question, while students who scored lower on the subscale were less likely to do so. In interpreting the item map information, it should be kept in mind that students at different grades demonstrated these abilities with grade-appropriate materials.

Figure 5.4 Map of Selected Items on the History Theme Subscales for Grade 4

Each U.S. history question was mapped onto the NAEP U.S. history theme subscale based on students' performance. The point on the subscale at which a question is positioned on the map represents the subscale score attained by students who had a 65-percent probability of successfully answering the question. Thus, it can be said for each question and its corresponding subscale score that students with scores above that point on the subscale have a greater than 65-percent chance of successfully answering the question, while those below that point on the subscale have a less than 65-percent chance. (The probability was set at 74 percent for multiple-choice questions.) In interpreting the item map information, it should be kept in mind that students at different grades demonstrated these abilities with grade-appropriate materials.



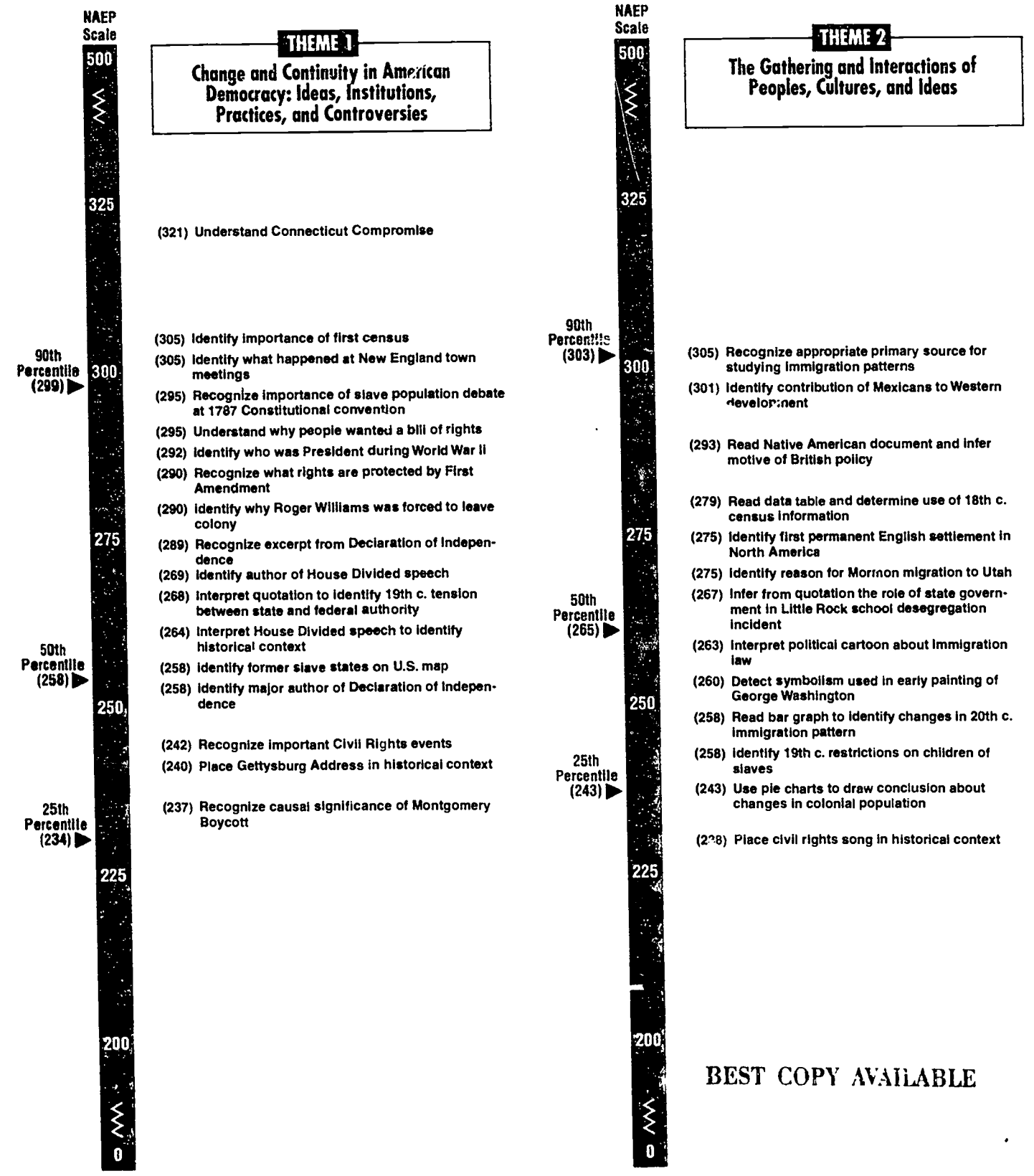
NOTE: In this graphic illustration, the locations of scale points are necessarily *approximate* for questions clustered closely together



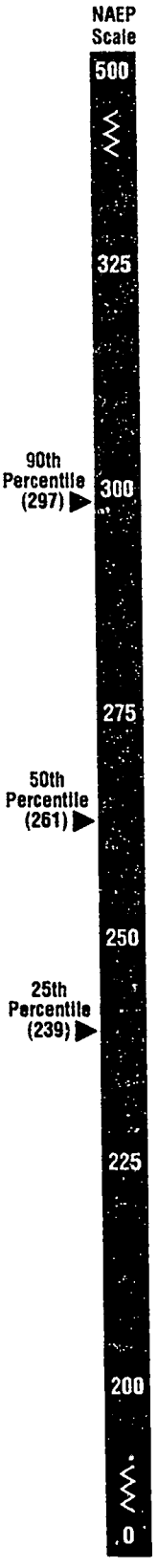
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Figure 5.5 Map of Selected Items on the History Theme Subscales for Grade 8

Each history question was mapped onto the NAEP history theme subscale based on students' performance. The point on the subscale at which a question is positioned on the map represents the subscale score attained by students who had a 65-percent probability of successfully answering the question. Thus, it can be said for each question and its corresponding subscale score that students with scores above that point on the subscale have a greater than 65-percent chance of successfully answering the question, while those below that point on the subscale have a less than 65-percent chance. (The probability was set at 74 percent for multiple-choice questions.) In interpreting the item map information, it should be kept in mind that students at different grades demonstrated these abilities with grade-appropriate materials.



NOTE: In this graphic illustration the locations of scale points are necessarily *approximate* for questions clustered closely together



THEME 3

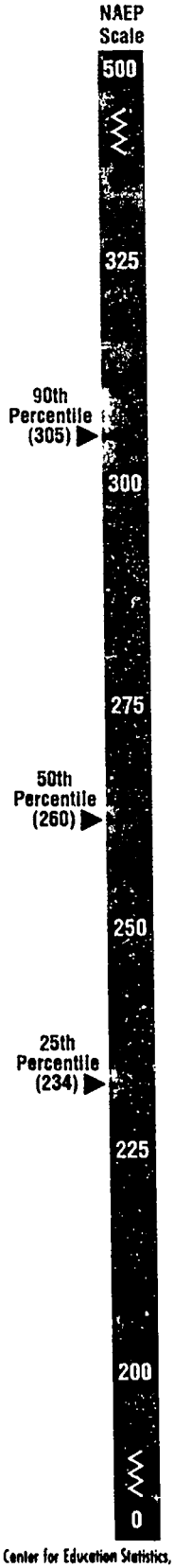
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

(297) Recognize movement inspired by Rachel Carson's *Silent Spring*
(296) Identify reason for New Deal
(294) Use population graph to describe changes in U.S. life between 1790 and 1970
(292) Recognize main effect of cotton gin
(290) Understand three useful sources for report on work during American Revolution
(286) Describe effects on West of building of railroad
(278) Use map to infer purpose of water supply in 19th c. town
(276) Use map to identify colonial trade route
(269) Recognize U.S. produced product from 19th c. grocery list

(251) Identify union strategy against factory owners
(250) Recognize effects of introduction of barbed wire
(247) Identify which invention most changed work life

(237) Recognize consequence of Gold Rush

(229) Infer report topic relevant to 19th c. social protest photographs



THEME 4

The Changing Role of America in the World

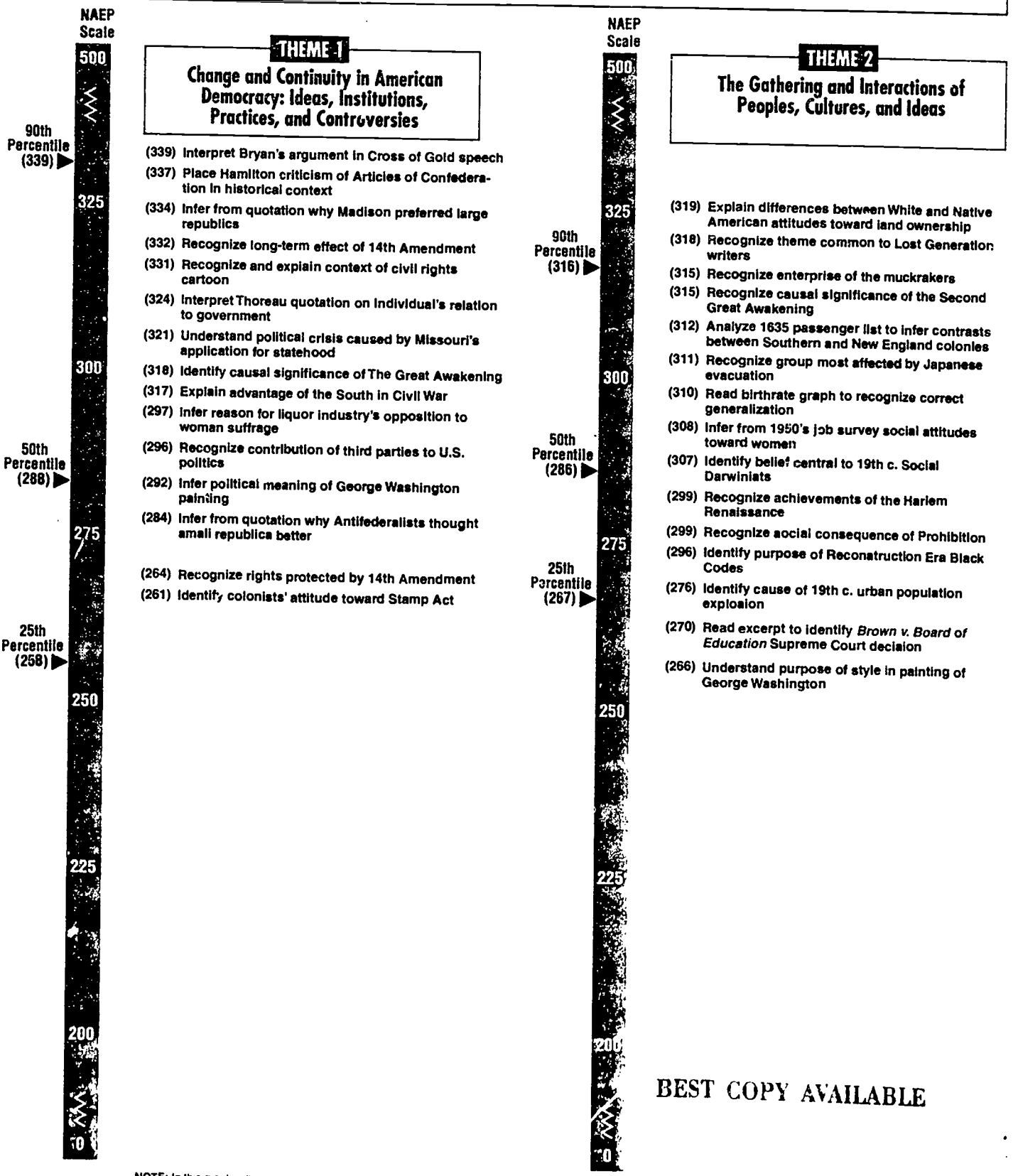
(312) Recognize events of World War II
(308) Describe important occurrence in 1776 Philadelphia
(305) Create time line of important colonial events
(305) Describe important colonial advantage in American Revolution
(297) Analyze 19th c. grocery list to identify U.S. foreign trade pattern
(295) Use world map to explain reasons for sending U.S. troops to two locations
(283) Recognize provision of 1783 Treaty of Paris
(283) Recognize which American Indian tribe lived in 15th c. Mexico
(281) Identify from which country U.S. acquired Florida
(279) Identify where U.S. dropped World War II atomic bombs
(271) Describe important British advantage in American Revolution
(260) Place Pearl Harbor poster in historical context

(255) Interpret purpose of 1943 World War II poster

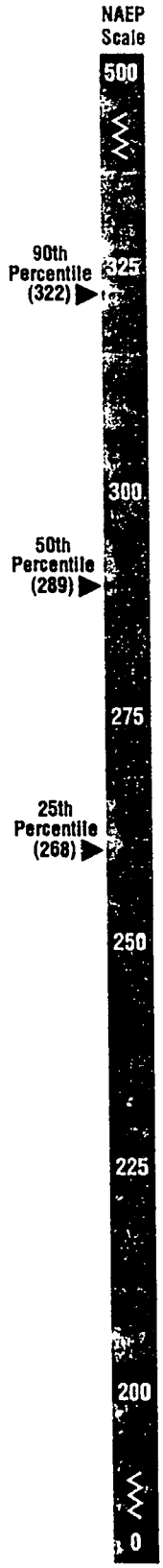
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Figure 5.6 Map of Selected Items on the History Theme Subscales for Grade 12

Each history question was mapped onto the NAEP history theme subscale based on students' performance. The point on the subscale at which a question is positioned on the map represents the subscale score attained by students who had a 65-percent probability of successfully answering the question. Thus, it can be said for each question and its corresponding subscale score that students with scores above that point on the subscale have a greater than 65-percent chance of successfully answering the question, while those below that point on the subscale have a less than 65-percent chance. (The probability was set at 74 percent for multiple-choice questions.) In interpreting the item map information, it should be kept in mind that students at different grades demonstrated these abilities with grade-appropriate materials.



NOTE: In this graphic illustration, the locations of scale points are necessarily *approximate* for questions clustered closely together.



THEME 3

Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

(334) Explain factor in changed nature of farming

(333) Compare F.D. Roosevelt's 1933 and 1937 inauguration speeches

(315) Explain factor in suburban growth

(314) Identify Upton Sinclair's book critiquing meat-packing industry

(311) Infer Civil War damage to urban South from news report

(310) Identify 19th c. supporters of gold standard

(309) Use map to identify important product from colonial trade route

(303) Recognize views of New Deal supporters

(300) Recognize importance of development of mail-order business

(298) Place quotation about 19th c. textile workers in historical context

(291) Analyze document to identify cause of pre-Depression prosperity

(290) Identify Virginia's 18th c. main cash crop

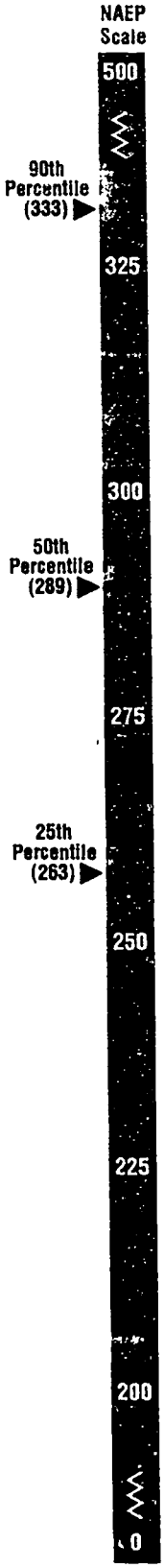
(286) Interpret cartoon to identify U.S. response to Soviet space exploration

(282) Identify motivation for European voyages of exploration

(277) Interpret photograph of women workers to identify pro-union protest

(275) Identify affect of small pox and measles vaccinations

(258) Interpret documents to identify F.D. Roosevelt's New Deal goals



THEME 4

The Changing Role of America in the World

(333) Explain John F. Kennedy quotation relating China and Vietnam

(332) Recognize purpose of Monroe Doctrine

(331) Infer information about presidential election from Vietnam graph

(329) Recognize motive in U.S. support of Marshall plan

(327) Analyze Woodrow Wilson quote about post-World War I peacekeeping

(323) Interpret message of Theodore Roosevelt foreign policy cartoon

(321) Identify factor in U.S. World War I involvement

(320) Recognize U.S. foreign policy goal, 1945-1990

(320) Analyze target of Harding campaign poster

(318) Interpret and place Soviet policy cartoon in historical context

(314) Recognize reason for British support of Confederacy

(313) Analyze message of avenge Pearl Harbor poster

(313) Identify energy crisis as threat to U.S. economy

(310) Interpret anti-League of Nations cartoon

(304) Recognize Cuban Missile Crisis as source of U.S./Soviet tension

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Summary

Examining specific knowledge and skills demonstrated by students at various points on the composite scale reveals a variety of patterns in student performance. Generally, students across grades in the higher percentiles exhibited greater abilities to analyze, draw inferences, work with primary documents, and think critically. Perhaps predictably, these students were more likely to read more pages per day in school, spend more time on their homework, and discuss their studies at home more frequently than their lower-performing peers. (See endnote 1.)

Illustrating student performance on specific tasks on item maps for each subscale helped place the assessment results in perspective. Clearly, students found the assessment challenging. For each subscale at grades 4, 8, and 12, the majority of tasks mapped above the 50th percentile, in many cases exceeding the 90th percentile. This indicates that many students had a less than 65-percent chance of performing well on these questions. The NAEP 1994 U.S. history assessment was rigorous; many tasks demanded knowledge of complex events and concepts, and abilities to analyze and interpret.

Patterns of student subgroup performance on the four subscales for the most part reflected those evidenced on the composite scale. For example, across the four subscales, as on the composite scale, White and Asian students had higher average scale scores than their Black and Hispanic peers. Also, there was a positive relationship between levels of parental education and student scores across subscales. However, there were some interesting variations in subscale performance as compared to that on the composite scale, the most notable being that, at grades 8 and 12, females outperformed males on the People, Cultures, and Ideas theme.

Endnotes

1. It should be noted that results for pages read in school and time spent on homework are based on collapsed data. The data do not necessarily imply a direct, positive linear relationship between student performance on the NAEP U.S. history assessment, and students' reports of pages read in school and time spent on homework.

CONCLUSION

The NAEP 1994 U.S. history assessment was constructed to be challenging. It required students to marshal bodies of complex knowledge, as well as to show their abilities to analyze, to explore points of view, and to think critically about U.S. history. A variety of stimuli, such as photographs, cartoons, and primary documents, were used. Also, a large percentage of assessment time was devoted to constructed-response questions for which students had to write their own answers.

Student performance on the assessment reflected its challenging nature. Many students found the assessment difficult. For the nation as a whole, few students reached the *Proficient* achievement level — defined as signifying solid grade-level performance — and only 1 or 2 percent reached the *Advanced* achievement level. This shows a much lower level of attainment than has been seen in other subjects assessed by NAEP. Moreover, when student performance on specific tasks is illustrated on item maps for each NAEP thematic subscale, many tasks fell well above the 50th percentiles for each subscale, indicating that only top-performing students could successfully respond to these questions on a consistent basis.

Some of the patterns in performance, however, are characteristic of those seen in other NAEP assessments, such as reading and geography. For example, White and Asian students generally had higher average scale scores than did Black and Hispanic students. Also, as seen in other NAEP assessments, students who reported that their parents had higher levels of education outperformed their peers who reported lower levels of parental education, and students attending nonpublic schools outperformed those in public schools.

Male students outperformed female students only at grade 12 on the U.S. history composite scale. However, subscale performances did not necessarily mirror performances on the composite scale. For example, female students at grades 8 and 12 outperformed males on the People, Cultures, and Ideas subscale.

Performance associated with home support and other background variables also was consistent with that found in other NAEP assessments. Frequent television watching was associated with lower scale scores, while having literacy materials and discussing studies at home were associated with higher scores. It is encouraging to note that exposure to U.S. history is associated with higher scores for grades 4 and 8 students, and the more semesters of potential history-related coursework done by twelfth graders, the better their performance.

While the NAEP results presented in this report cannot be used to draw causal inferences, they nevertheless do point out interesting characteristics and patterns of student performance. Future research and other projects and analyses can utilize NAEP data to shed more light on relationships between performance and background data, that in turn can be used by policymakers, educators, and citizens to bring change to the United States educational system.

APPENDIX A

Overview of Procedures Used in NAEP's 1994 U.S. History Assessment

Introduction

The conduct of a large-scale assessment of educational progress entails the successful coordination of a multitude of projects, committees, procedures, and tasks. This appendix provides an overview of the NAEP 1994 U.S. history assessment's primary components — framework, development, administration, scoring, and analysis. A more extensive review of procedures and methods utilized in the assessment will be included in a subsequent technical report: *The NAEP 1994 Technical Report*.

NAEP's U.S. History Assessment Framework

The framework underlying the NAEP 1994 U.S. history assessment reflects current consensus among educators and researchers about the study of U.S. history.

The framework's purpose was to provide a view of U.S. history on which to base the NAEP assessment. Developing this framework and the specifications that guided development of the assessment involved the critical input of hundreds of individuals across the country, including representatives of national education organizations, teachers, parents, policymakers, business leaders, and the interested general public. This consensus process was managed by the Council of Chief State School Officers for the National Assessment Governing Board.

The framework sets forth a broad content matrix that describes U.S. history in terms of historical themes and periods. Figure A.1 summarizes this content matrix.

Figure A.1 NAEP 1994 U.S. History Content Matrix

THEMES / PERIODS	Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	The Gathering and Interactions of Peoples, Cultures, and Ideas	Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	The Changing Role of America in the World
Three Worlds and Their Meeting in the Americas (Beginnings to 1607)				
Colonization, Settlement, and Communities (1607 to 1763)				
The Revolution and the New Nation (1763 to 1815)				
Expansion and Reform (1801 to 1861)				
Crisis of the Union: Civil War and Reconstruction (1850 to 1877)				
The Development of Modern America (1865 to 1920)				
Modern America and the World Wars (1914 to 1945)				
Contemporary America (1945 to Present)				

The assessment framework specified not only the particular aspects of U.S. history to be measured, but also the percentage of assessment time that should be devoted to each. The target percentage distributions of historical themes and historical periods as specified in the framework, along with the actual percentage distributions in the assessment, are presented in Tables A.1 and A.2. The actual content of the assessment was consistent with the targeted distribution.

TABLE A.1

Target and Actual Percentage Distribution of Questions by Grade and Historical Theme Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

Historical Themes	Grade 4		Grade 8		Grade 12	
	Target	Actual	Target	Actual	Target	Actual
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	25%	24%	30%	28%	25%	29%
The Gathering and Interactions of Peoples, Cultures, and Ideas	35%	32%	30%	30%	25%	23%
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	25%	25%	20%	23%	25%	26%
The Changing Role of America in the World	15%	19%	20%	19%	25%	21%

TABLE A.2

Target and Actual Percentage Distribution of Questions by Grade and Historical Period Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

Historical Periods	Grade 4		Grade 8		Grade 12	
	Target	Actual*	Target	Actual*	Target	Actual*
Three Worlds and Their Meeting in the Americas (Beginnings to 1607)	20%	15%	5%	3%	5%	2%
Colonization, Settlement, and Communities (1607 to 1763)	15%	13%	10%	7%	10%	8%
The Revolution and the New Nation (1763 to 1815)	15%	12%	20%	20%	15%	10%
Expansion and Reform (1801 to 1861)	15%	10%	15%	10%	10%	7%
Crisis of the Union: Civil War and Reconstruction (1850 to 1877)	10%	8%	20%	13%	10%	10%
The Development of Modern America (1865 to 1920)	5%	9%	10%	13%	15%	15%
Modern America and the World Wars (1914 to 1945)	5%	5%	10%	12%	15%	29%
Contemporary America (1945 to Present)	15%	15%	10%	11%	20%	20%

* Actual period percentages do not add up to 100 percent, due to the classification of some items in more than one period, or generally applied across historical periods

The Assessment Design

Each student in the assessment received an assessment booklet containing general background questions, U.S. history questions, a set of background questions specific to social studies or history, and a set of questions that determined students' motivation and familiarity with the assessment tasks. The U.S. history questions and their stimulus material were arranged into blocks. Students were given either two 25-minute blocks or one 50-minute block. At the fourth grade, only 25-minute blocks were used.

The grade 4 assessment consisted of six 25-minute blocks, two of which were also administered to eighth-grade students. Each block contained single questions and sets, a variety of stimulus material, and a combination of multiple-choice and constructed-response questions. One of the constructed-response questions in each block required an extended response. A total of 63 multiple-choice questions, 26 short constructed-response questions, and six extended constructed-response questions were administered at grade 4.

The grade 8 assessment consisted of eight 25-minute blocks (including two also administered at grade 4 and two also administered at grade 12), and one 50-minute block. The 25-minute blocks followed the same pattern as those for grade 4, except that some included more than one extended constructed-response question. The 50-minute block included questions all focusing on a particular theme, and contained three extended constructed-response questions. A total of 102 multiple-choice questions, 37 short constructed-response questions, and 12 extended constructed-response questions were administered at grade 8.

The grade 12 assessment consisted of eight 25-minute blocks (including two also administered at grade 12), and one 50-minute block. These blocks followed the same pattern as those for grades 4 and 8, except that all blocks included two extended constructed-response questions. The 50-minute block included questions all focusing on a particular theme, contained four extended constructed-response questions. A total of 103 multiple-choice questions, 33 short constructed-response questions, and 19 extended constructed-response questions were administered at grade 12.

The assessment design allowed for maximum coverage of the domain of U.S. history at each grade, while minimizing the time burden for any one student. This was accomplished through the use of matrix sampling, in which a representative sample of students takes each portion of the assessment. Individual students were required to take only a small part; however, the aggregate results across the entire assessment allow for broad reporting of U.S. history abilities for the targeted population.

In addition to matrix sampling, the assessment design used a procedure for distributing booklets that controlled for position and balance effects. Students received different blocks of questions in their booklets according to a specific design. Balanced incomplete block (BIB) spiraling was used to assign blocks of questions in a manner that balanced the positioning of blocks across booklets and balanced the pairing of blocks within booklets. The spiraling aspect of this procedure cycles the booklets for administration so that typically only a few students in any assessment session receive the same booklet.

Teacher and School Questionnaires

One of the most important parts of NAEP's efforts to document the nature of students' achievement is the collection of contextual information regarding students' school experiences. As a part of the 1994 U.S. history assessment, NAEP administered a questionnaire to teachers responsible for teaching social studies or history to students who participated in the fourth- or eighth-grade assessments. In addition, the principals or other administrators of sampled schools at all grades were asked to complete a school questionnaire. These questionnaires were developed in consultation with an expert panel. These instruments focused on five areas: instructional content, instructional practices and experiences, teacher characteristics, school conditions and contexts, and conditions outside the school (i.e., home support, out-of-school activities, and attitudes).

The fourth- and eighth-grade social studies and history teacher questionnaires were composed of two sections each. One section contained questions about teachers' background, education, and resources. Another section posed questions to teachers about their recent exposure to training in various areas of history education, the structure and nature of their classroom instruction, and the types of materials and approaches they use in teaching history.

Because the sampling of teachers for the teacher questionnaires was based on participating students, the teachers' questionnaire responses do not necessarily represent all fourth- and eighth-grade teachers in the nation. Rather, they represent teachers of the representative sample of students in the assessment.

Consequently, these findings portray the nature of students' instructional experiences and the background of their teachers.

It is important to note that in this report, as in all NAEP reports, the student is the unit of analysis — even when information from teacher or school questionnaires is being reported. Using the student as the unit of analysis makes it possible to link students' performance with their instructional and background experiences, thus providing a rich source of relevant information for educators and researchers. Although this approach may provide a different perspective from other studies that simply report information about teachers or schools, it is consistent with NAEP's goal of providing information about the educational context and performance of students.

Some students selected for the assessment were judged by school authorities to be incapable of meaningful participation in the assessment because they had limited English-language proficiency, were mentally challenged, or were functionally disabled. (See *Limited English Proficient and Individualized Education Plan* section in this appendix.) For each student excluded from the assessment, schools were required to complete a questionnaire about the characteristics of that student and the reason for exclusion.

NAEP U.S. History Samples

The results presented in this report are based on nationally representative probability samples of fourth-, eighth-, and twelfth-grade students. The samples were selected using a complex multistage sampling design involving the sampling of students from selected schools within selected geographic areas across the country. The sample design had the following stages:

1) selection of primary sampling units (PSUs) — geographic areas defined as counties or groups of counties);

2) selection of schools (both public and nonpublic) within the selected areas; and

3) selection of students within selected schools.

Each selected school that participated in the assessment, and each student assessed, represents a portion of the population of interest. To make valid inferences from the student samples to the respective populations from which they were drawn, sampling weights are needed. Sampling weights account for disproportionate representation due to oversampling of nonpublic schools and of students attending schools with high concentrations of Black and/or Hispanic students. Lower sampling rates for very small schools must also be accounted for with the sampling weights.

TABLE A.3

Unweighted and Weighted Sample Sizes by Grade Public and Nonpublic Schools Grades 4, 8, and 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Unweighted Sample Size (and Percent of Total)		
	Grade 4	Grade 8	Grade 12
Nation	5,499 (100.0%)	8,767 (100.0%)	7,818 (100.0%)
Region			
Northeast	1,334 (24.3%)	1,614 (18.4%)	1,770 (22.6%)
Southeast	1,440 (26.2%)	2,664 (30.4%)	2 168 (27.7%)
Central	1,241 (22.6%)	1,821 (20.8%)	1,575 (20.1%)
West	1,484 (27.0%)	2,668 (30.4%)	2,305 (29.5%)

	Weighted Sample Size (and Percent of Total)		
	Grade 4	Grade 8	Grade 12
Nation	3,527,794 (100.0%)	3,449,193 (100.0%)	2,545,898 (100.0%)
Region			
Northeast	767,788 (21.8%)	692,421 (20.1%)	517,336 (20.3%)
Southeast	818,785 (23.2%)	878,674 (25.5%)	582,651 (22.9%)
Central	887,761 (25.2%)	820,691 (23.8%)	683,496 (26.8%)
West	1,053,461 (29.9%)	1,057,406 (30.7%)	762,415 (29.9%)

Percentages may not total 100 percent due to rounding

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment

Table A.3 provides a summary of the weighted and unweighted student sample sizes for the U.S. history assessment. The numbers reported include both public and nonpublic school students.

Limited English Proficiency (LEP) and Individualized Education Plan (IEP) Students

It is NAEP's intent to assess all selected students. Therefore, every effort is made to ensure that all selected students who are capable of participating in the assessment are assessed. However, some students sampled for participation in NAEP can be excused from the sample according to carefully defined criteria. Specifically, some students identified as having Limited English Proficiency (LEP) or having an Individualized Education Plan (IEP) may be incapable of participating meaningfully in the assessment. These students are identified as follows:

Students classified as LEP may be excluded from the assessment if

- the student is a native speaker of a language other than English; AND
- the student has been enrolled in an English-speaking school less than two years; AND
- the school staff most familiar with the student have judged the student to be incapable of taking part in the assessment.

Students classified as IEP may be excluded if

- the student is mainstreamed less than 50 percent of the time in academic subjects and is judged to be incapable of taking part in the assessment, OR
- the IEP team has determined that the student is incapable of taking part meaningfully in the assessment.

When there is doubt, the student is included in the assessment.

For each student excused from the assessment, school personnel complete a questionnaire about the characteristics of that student and the reason for exclusion.

Data Collection

The NAEP 1994 U.S. history assessment was conducted from January through March 1994, with some makeup sessions in early April. As with all NAEP assessments, data collection for the 1994 assessment was conducted by trained field staff. For the U.S. history assessment, this was accomplished by Westat, Inc., staff.

Scoring

Materials from the 1994 assessment were shipped to National Computer Systems in Iowa City for processing. Receipt and quality control were managed through a sophisticated bar-coding and tracking system. After all appropriate materials were received from a school, they were forwarded to the professional scoring area where the responses to the constructed-response questions were evaluated by trained staff using guidelines prepared by NAEP. Each constructed-response question had a unique scoring guide that defined the criteria to be used in evaluating students' responses. The extended constructed-response questions were evaluated with four-level rubrics, and many of the short constructed-response questions were rated according to three-level rubrics that permitted partial credit to be given.

For the NAEP 1994 U.S. history assessment, approximately 375,000 student responses were scored. This figure includes a 25 percent rescore to monitor interrater reliability. The overall percentages of agreement between scorers for the 1994 reliability samples were 90 percent at grade 4, 90 percent at grade 8, and 89 percent at grade 12.

Data Analysis and IRT Scaling

Subsequent to the professional scoring, all information was transcribed to the NAEP database at ETS. Each processing activity was conducted with rigorous quality control. After the assessment information had been compiled in the database, the data were weighted according to the population structure. The weighting for the samples reflected the probability of selection for each student as a result of the sampling design, adjusted for nonresponse. Through stratification, the weighting assured that the representation of certain subpopulations corresponded to figures from the U.S. Bureau of the Census and the Current Population Survey.[1]

Analyses were then conducted to determine the percentages of students who gave various responses to each U.S. history and background question. In determining these percentages for the U.S. history questions, a distinction was made between missing responses at the end of a block (i.e., missing responses subsequent to the last question the student answered) and missing responses prior to the last observed response. Missing responses before the last observed response were considered intentional omissions. Missing responses at the end of the block were considered "not reached" and treated as if the questions had not been presented to the student. In calculating response percentages for each question, only students classified as having been presented the question were included in the denominator of the statistic.

It is standard ETS practice to treat all nonrespondents to the last question in a block as if they had not reached the question. For multiple-choice and short constructed-response questions, this practice produces a reasonable pattern of results in that the proportion reaching the last question is not dramatically smaller than the proportion reaching the next-to-last question. However, for blocks that ended with extended constructed-response questions, the standard ETS practice would result in extremely large drops in the proportion of students attempting the final question. A drop of such magnitude seemed somewhat implausible. Therefore, for blocks ending with an extended constructed-response question, students who answered the next-to-last question but did not respond to the extended constructed-response question were classified as having intentionally omitted the last question.

Item response theory (IRT) was used to estimate average U.S. history scale scores for the nation and for various subgroups of interest within the nation. IRT models the probability of answering a question in a certain way as a mathematical function of skill. The main purpose of IRT analysis is to provide a common scale on which performance can be compared across groups such as those defined by grades and characteristics, including race/ethnicity and gender.

Because of the BIB-spiraling design used by NAEP, students do not receive enough questions about a specific topic to provide reliable information about individual performance. Traditional test scores for individual students, even those based on IRT, would lead to misleading estimates of population characteristics, such as subgroup means and percentages of students at or above a certain achievement level. Consequently, NAEP constructs sets of plausible values designed to represent the distribution of scores in the population. A plausible value for an individual is not a scale score for that individual but may be regarded as a representative value from the distribution of potential scale scores for all students in the population with similar characteristics and identical patterns of item response. Statistics describing performance on the NAEP U.S. history scale are based on the plausible values. They estimate values that would have been obtained had individual performances been observed — that is, had each student responded to a sufficient number of cognitive questions so that performance could be precisely estimated.[2]

For the NAEP 1994 U.S. history assessment, within-grade scales were created to report performance for each subscale (i.e., historical theme). Similar scaling procedures were used to establish each of the four subscales. Specifically, three within-grade subscales (one for each grade) were established for each of the U.S. history content areas. The within-grade subscales for grades 4 and 12 were then linked to the grade 8 subscale to form a common reporting metric. This common reporting metric, which runs from 0 to 500 for each of the subscales, was established so that the mean scores across all three grades is 250 and the standard deviation of the scores is 50.

The composite NAEP U.S. history scale was produced as a weighted average of the subscales, the weights being given by the target percentages shown in Table A.1. The reporting metric of the composite scale, which also runs from 0 to 500, was again established so that the mean score across all three grades is 250. No constraints were imposed on the standard deviation of the cross-grade composite scores.

It may be helpful here to provide some guidance to the reader of this report about the types of cross-grade and cross-scale inferences that are appropriate. The use of a common cross-grade metric for the subscales and the composite scale was motivated primarily by issues of convenience in the reporting of results. In producing the subscales, IRT parameters for questions common across the grades were not constrained to be equal. Furthermore, as is evident in Table A.1, the weights used to combine the subscales into the composite U.S. history scale differ by grade. As a result of these two scaling conventions, cross-grade comparisons of scale score averages, both at the subscale level as well as for the composite scale, may not be meaningful. Similarly, scale score differences (e.g., between subscale or

composite scale averages for males and females) should probably not be compared across grades. The reader is best served by focusing on within-grade group comparisons and inferences.

The use of a common scaling procedure for each of the subscales does provide for some within-grade normative meanings across subscales. For example, at grade 4, a score of 200 is one cross-grade standard deviation unit below the cross-grade average for each of the U.S. history subscales. Similarly, a score of 210 is four-fifths of a cross-grade standard deviation below the cross-grade average for each of the scales. A group, for example males, scoring 200 on the American Democracy subscale and 210 on the Economic and Technological Changes subscale did indeed perform "better" on the latter than on the former *in a cross-grade normative sense*. However, other inferences about relative performance in, say, a percent-correct metric do not necessarily follow. For example, a score of 200 on the American Democracy subscale may imply a higher expected percent-correct score on the collection of assessment exercises that define that scale than is implied by a score of 210 on the Economic and Technological Changes subscale. Thus, continuing with the current example, performance on the American Democracy subscale was better *in the percent-correct sense* than performance on the Economic and Technological Changes scale.

In Chapter 5, performance across subscales is compared by examining patterns of subgroup differences (i.e., patterns of statistical significance between subgroups across the four subscales). These patterns are discussed separately for each of the three grades. Within-grade inferences based on such comparisons are defensible given the limited degree of comparability that exists in the subscale reporting metrics. As noted above, other types of inferences (e.g., inferences involving subscale score differences) may be less defensible.

The subscales summarize student performance across all three question types in the assessment (multiple-choice, short constructed-response, and extended constructed-response). In producing these subscales, two IRT models were used. Multiple-choice questions were scaled using the three-parameter logistic (3PL) model; and short constructed-response questions rated according to a three-level rubric, as well as extended constructed-response questions rated on a four-level rubric, were scaled using a generalized partial-credit (GPC) model.[3] Developed by ETS and first used in 1992, the GPC model permits the scaling of questions scored according to multipoint rating schemes. The model takes full advantage of the information available from each of the student response categories used for these more complex constructed-response questions.

The U.S. history scale is composed of two types of questions: multiple-choice and constructed-response (scored according to a partial-credit model). One natural question about the scale concerns the amount of information contributed by each type of question. Unfortunately, this question has no simple answer for the NAEP U.S. history assessment, due to the complex procedures used to form the composite U.S. history scale.

The information provided by a given question is determined by the IRT model used to scale the question and is a function of proficiencies.[4] Thus, the answer to the query "How much information do the different types of questions provide?" will differ for each level of U.S. history proficiency. When considering the composite U.S. history scale, the answer is even more complicated. The U.S. history data are scaled separately by the historical themes. As discussed on the previous page, the composite scale is a weighted combination of these subscales. IRT information functions are only strictly comparable when they are linked on a common scale. Because the composite scale is based on four separate calibrations, without any common item-linking, there is no direct way to compare the information provided by the questions on the composite scale.

NAEP Reporting Groups

Findings from the NAEP 1994 U.S. history assessment are presented for groups of students defined by shared characteristics. Data are reported for subgroups only where sufficient numbers of students and adequate school representation are present. There must be at least 62 students in a particular subgroup, and these students must come from at least six different PSUs (see description of sampling design on page 80). Data for all students, regardless of whether their subgroup was reported separately, were included in computing overall national and regional results.

The reporting subgroups presented in this report include: region, race/ethnicity, gender, parents' education level, type of school, and school's type of location. Definitions of these subgroups are provided on the following page.

Region. Results are reported for four regions of the nation: Northeast, Southeast, Central and West. States included in each region are shown in Figure A.2. All 50 states and the District of Columbia are listed. U.S. territories were not assigned to a region.

Race/Ethnicity. The race/ethnicity variable is an imputed definition of race/ethnicity, derived from up to three sources of information. This variable is used for race/ethnicity subgroup comparisons. Two items from the student demographics questionnaire were used in the determination of derived race/ethnicity:

Demographic Item Number 2:

2. If you are Hispanic, what is your Hispanic background?

   ○ I am not Hispanic.

   ○ Mexican, Mexican American, or Chicano

   ○ Puerto Rican

   ○ Cuban

   ○ Other Spanish or Hispanic background

Students who responded to Item Number 2 by filling in the second, third, fourth, or fifth oval were considered Hispanic. For students who filled in the first oval, did not respond to the item, or provided information that was illegible or could not be classified, responses to Item Number 1 were examined in an effort to determine race/ethnicity. Item Number 1 read as follows:

Demographic Item Number 1:

1. Which best describes you?

   ○ White (not Hispanic)

   ○ Black (not Hispanic)

   ○ Hispanic ("Hispanic" means someone who is Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or other Spanish or Hispanic background.)

   ○ Asian ("Asian" means someone who is Chinese, Japanese, Korean, Vietnamese, or other Asian background.)

   ○ Pacific Islander ("Pacific Islander" means someone who is from a Filipino, Hawaiian, or other Pacific Island background.)

   ○ American Indian or Alaskan Native ("American Indian or Alaskan Native" means someone who is from one of the American Indian tribes, or one of the original people of Alaska.)

   ○ Other

Students' race/ethnicity was then assigned to correspond with their selection. For students who filled in the seventh oval ("Other"), provided illegible information or information that could not be classified, or did not respond at all, race/ethnicity as provided from school records was used.

Derived race/ethnicity could not be determined for students who did not respond to demographic items 1 or 2 and for whom race/ethnicity was not provided by the school.

Figure A.2 States Included in the Four Regions

NORTHEAST	SOUTHEAST	CENTRAL	WEST
Connecticut	Alabama	Illinois	Alaska
Delaware	Arkansas	Indiana	Arizona
District of Columbia	Florida	Iowa	California
Maine	Georgia	Kansas	Colorado
Maryland	Kentucky	Michigan	Hawaii
Massachusetts	Louisiana	Minnesota	Idaho
New Hampshire	Mississippi	Missouri	Montana
New Jersey	North Carolina	Nebraska	Nevada
New York	South Carolina	North Dakota	New Mexico
Pennsylvania	Tennessee	Ohio	Oklahoma
Rhode Island	Virginia*	South Dakota	Oregon
Vermont	West Virginia	Wisconsin	Texas
Virginia*			Utah
			Washington
			Wyoming

*Note that the part of Virginia that is included in the Washington, DC, metropolitan area is included in the Northeast region; the remainder of the state is included in the Southeast region.

Gender. Results are reported separately for males and females.

Parents' Education Level. The parents' education level variable is derived from responses to two questions in the student demographic questionnaire. Students were asked to indicate the extent of their mother's education (How far in school did your mother go?) by choosing one of the following:

- ○ She did not finish high school.
- ○ She graduated from high school.
- ○ She had some education after high school.
- ○ She graduated from college.
- ○ I don't know.

Students were asked to provide the same information about the extent of their father's education (How far in school did your father go?) by choosing one of the following:

- ○ He did not finish high school.
- ○ He graduated from high school.
- ○ He had some education after high school.
- ○ He graduated from college.
- ○ I don't know.

The information was combined into one parental education reporting category as follows: if a student indicated the extent of education for only one parent, that level was included in the data. If a student indicated the extent of education for both parents, the higher of the two levels was included in the data. For students who did not know the level of education for both parents or did not know the level of education for one parent and did not respond for the other, the parental education level was classified as unknown. If the student did not respond for both parents, the student was recorded as having provided no response.

It should be noted that approximately one-third of fourth graders and almost one-tenth of eighth graders reported not knowing the education level of either of their parents. The percentages of students who reported not knowing their parents' education level were larger for fourth-grade Hispanic students and for eighth-grade Black and Hispanic students compared to their White counterparts. (See Table A.4.)

TABLE A.4

Percentage of Students Who Reported Not Knowing Their Parents' Education Level, by Race/Ethnicity Grades 4, 8, and 12 1994 U.S. History Assessment

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	Total	White	Black	Hispanic
Grade 4	34 (0.8)	32 (1.0)	31 (1.3)	43 (2.1)
Grade 8	9 (0.4)	7 (0.5)	11 (0.8)	20 (1.5)
Grade 12	3 (0.2)	2 (0.2)	5 (0.8)	8 (0.8)

The standard errors of the percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

In addition, evidence from other NCES surveys that gather data from students and parents indicates larger discrepancies between students' and parents' reports for Black and Hispanic students compared to White students. These differences between racial/ethnic groups are more evident at grade 8 than grade 12. As shown in Table A.5, the correlations between students' and parents' reports of parental education were lower for Black and Hispanic students than for White students at both grades 8 and 12, although all correlations were higher in twelfth grade.

TABLE A.5

Correlations Between Students' and Parents' Reports of Parents' Education Level, by Race/Ethnicity* Grades 8 and 12

	White	Black	Hispanic
Grade 8			
Father's Education	0.84	0.67	0.75
Mother's Education	0.79	0.62	0.65
Grade 12			
Father's Education	0.90	0.80	0.85
Mother's Education	0.87	0.78	0.74

* These results are from NCES surveys other than NAEP.

SOURCE: **For grade 8** – P. Kaufman and R.A. Rosinski, *Quality of Responses of Eighth-Grade Students in NELS: 88*, Washington, DC: National Center for Education Statistics, NCES 91-487; **For grade 12** – W. F. Fetters, P.S. Stowe, and J.A. Owings, *Quality of Responses of High School Students to Questionnaire Items*, Washington, DC: National Center for Education Statistics, NCES 84-342.

Type of School. Results are reported by the type of school that the student attends: public or nonpublic. Nonpublic schools include Catholic and other nonpublic schools. Bureau of Indian Affairs (BIA) schools and domestic Department of Defense (DoD) schools are not included in either the public or nonpublic categories but are included in the overall national results.

Type of Location. Results are reported for students attending schools in three mutually exclusive location types: central city, urban fringe/large town, and rural/small town:

Central City: This category includes central cities of all Standard Metropolitan Statistical Areas (SMSAs).[5] Central City is a geographic term and is not synonymous with "inner city."

Urban Fringe/Large Town: The urban fringe category includes all densely settled places and areas within SMSAs that are classified as urban by the U.S. Bureau of the Census. A Large Town is defined as a place outside a SMSA with a population greater than or equal to 25,000.

Rural/Small Town: Rural includes all places and areas with populations of less than 2,500 that are classified as rural by the U.S. Bureau of the Census. A Small Town is defined as a place outside a SMSA with a population of less than 25,000 but greater than or equal to 2,500.

As described earlier, the NAEP U.S. history scale makes it possible to examine relationships between students' performance and a variety of background factors measured by NAEP. However, the fact that a relationship exists between achievement and another variable does not reveal the underlying cause of the relationship, which may be influenced by a number of other variables. Similarly, the assessment does not capture the influence of unmeasured variables. The results are most useful when they are considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

Estimating Variability

Because the statistics presented in this report are estimates of group and subgroup performance based on samples of students, rather than the values that could be calculated if every student in the nation answered every question, it is important to account for the degree of uncertainty associated with the estimates. Two components of uncertainty are accounted for in the variability of statistics based on scale scores: 1) the uncertainty due to sampling only a relatively small number of students, and 2) the uncertainty due to sampling only a relatively small number of questions. The variability associated with the estimated percentages of students with certain background characteristics or who answered a certain cognitive question correctly is accounted for by the first component alone.

In addition to providing estimates of percentages of students and their average scores, this report provides information about the uncertainty of each statistic. Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate. NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any information about students that can be observed without error. However, each student typically responds to so few questions within any content area that the score measurement for any single student would be imprecise. In this case, using plausible-values technology makes it possible to describe the performance of groups and subgroups of students, but the underlying imprecision that makes this step necessary adds an additional component of variability to statistics based on NAEP scale scores.[6]

The reader is reminded that, like findings from all surveys, NAEP results are also subject to other kinds of errors including the effects of imperfect adjustment for student and school nonresponse, and other unknowable effects associated with the particular instrumentation and data collection methods. Nonsampling errors can be attributed to a number of sources: inability to obtain complete information about all selected schools in the sample (some students or schools refused to participate, or students participated but answered only certain questions); ambiguous definitions; differences in interpreting questions; inability or unwillingness to give correct information; mistakes in recording, coding, or

scoring data; and other errors of collecting, processing, sampling, and estimating missing data. The extent of nonsampling error is difficult to estimate. By their nature, the impact of such errors cannot be reflected in the data-based estimates of uncertainty provided in NAEP reports.

Drawing Inferences from the Results

The use of *confidence intervals*, based on the standard errors, provides a way to make inferences about the population means and percentages in a manner that reflects the uncertainty associated with the sample estimates.

An estimated sample mean ± 2 standard errors represents a 95-percent confidence interval for the corresponding population quantity. This means that with approximately 95-percent certainty, the average scale score for the entire population of interest is within ± 2 standard errors of the sample mean.

As an example, suppose that the average score of students in a particular group was 256, with a standard error of 1.2. A 95-percent confidence interval for the population quantity would be as follows:

Mean ± 2 standard errors
$256 \pm 2 \times 1.2$

256 ± 2.4

253.6, 258.4

Thus, one can conclude with 95-percent certainty that the average scale score for the entire population of students in that group is between 253.6 and 258.4.

Similar confidence intervals can be constructed for percentages, provided that the percentages are not extremely large (greater than 90) or extremely small (less than 10). For extreme percentages, confidence intervals constructed in the above manner may not be appropriate. However, procedures for obtaining accurate confidence intervals are quite complicated. Thus, comparisons involving extreme percentages should be interpreted with this in mind.

To determine whether there is a real difference between the mean score (or percentage of a certain attribute) for two groups in the population, one needs to obtain an estimate of the degree of uncertainty associated with the difference between the means or percentages of these groups for the sample. When comparing two independent estimates, this estimate of the degree of uncertainty — called the standard error of the difference between the groups — is obtained by taking the square of each group's standard error, summing these squared standard errors, and then taking the square root of this sum.

$$SE_{AB} = \sqrt{SE_A^2 + SE_B^2}$$

Similar to the manner in which the standard error for an individual group mean or percentage is used, the standard error of the difference can be used to help determine whether differences between groups in the population are real. The difference between the mean scale score or percentage of the two groups ± 2 standard errors of the difference represents an approximate 95-percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between groups in the population. If the interval does not contain zero, the difference between groups is statistically significant (different) at the .05 level.

The procedures described in this section, and the certainty ascribed to intervals (e.g., a 95-percent confidence interval) are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. When one considers sets of confidence intervals, statistical theory indicates that the certainty associated with the entire set of intervals is less than that attributable to each individual comparison from the set. If one wants to hold the certainty level for a specific set of comparisons at a particular level (e.g., 95), adjustments (called multiple-comparisons procedures) need to be made. A more complete discussion of the multiple-comparisons procedures is presented in the *NAEP 1994 Technical Report.*

The standard errors for means and percentages reported by NAEP are statistics and subject to a certain degree of uncertainty. In certain cases, typically when the standard error is based on a small number of students (or when the group of students is enrolled in a small number of schools), the amount of uncertainty associated with the standard errors may be quite large. Throughout this report, estimates of standard errors subject to a large degree of uncertainty are designated by the symbol "!". In such cases, the standard errors — and any confidence intervals or significance tests involving these standard errors — should be interpreted cautiously.

Endnotes

1. For additional information about the use of weighting procedures in NAEP, see Johnson, E. G. (1989). Considerations and techniques for the analysis of NAEP data. *Journal of Educational Statistics, 14*(4), 303-334.

2. For theoretical and empirical justification of the procedures employed, see Mislevy, R. J. (1988). Randomization-based inferences about latent variables from complex samples. *Pyschometrika, 56*(2), 177-196.

   For computational details, see Mislevy, R.J. (1990). Scaling procedures. In E.G. Johnson & R. Zwick, in collaboration with N. Allen, et al. *Focusing the new design: The NAEP 1988 technical report.* Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service. See also Mislevy, R.J. (1992). Scaling procedures. In E.G. Johnson & N. Allen, *The 1990 NAEP technical report.* Washington, DC: National Center for Education Statistics.

3. Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159-176.

4. Donoghue, J. R. (1994). An empirical examination of the IRT information of polytomously scored reading items under the generalized partial credit model. *Journal of Educational Measurement, 31*(4), 295-311.

   Muraki, E. (1993). Information functions of the generalized partial credit model. *Applied Psychological Measurement, 17*(4), 351-363.

5. Standard Metropolitan Statistical Area (SMSA) as defined by the Office of Management and Budget.

6. For further details, see Johnson, E.G., & Rust, K. F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics, 17*(2), 175-190.

APPENDIX B

Describing Students' U.S. History Performance

This appendix contains detailed information about the procedures used for describing students' U.S. history knowledge and abilities and profiling students' study habits. The results of these procedures are presented in Chapter 5 of this report.

Performance Descriptions Based on the U.S. History Composite Scale

A procedure known as scale anchoring was used to develop descriptions of student performance at selected points on the NAEP U.S. history composite scale. The scale points that were selected for anchoring reflect three levels of history knowledge and abilities corresponding to lower-, middle-, and higher-performing students. These levels correspond to the 25th, 50th, and 90th percentile points on the composite scale as established by the performance of students in 1994 — the first assessment administered under NAEP's current U.S. history framework.

Around each percentile point, a band was built to define a range of scale scores. Students described as being at a particular level were within a five percentile point range on either side of the specified scale point. For example, the 50th percentile was defined as the region between the 45th and 55th percentile points on the scale. A question was identified as anchoring at a percentile point on the scale if it was answered successfully by at least 65 percent of the students within that percentile band. (The criterion was set at 74 percent for multiple-choice questions to correct for the possibility of answering correctly by guessing.)

After defining the bands of the scale to be anchored, the next step in the process was to identify: (1) questions answered correctly for dichotomously scored questions, or (2) questions answered at a particular score level for partial credit constructed-response questions. Because the extended constructed-response questions were scored according to four levels of performance, each extended constructed-response question was treated as three distinct questions corresponding to scores of Partial or better, Essential or better, and Extensive. These distinct score levels were then analyzed in the same manner as questions scored dichotomously, as either correct or incorrect. Thus, for example, an extended constructed-response question might anchor at the 50th percentile for Partial or better responses **and** at the 90th percentile for Essential or better responses.

A committee of U.S. history education experts, including teachers for the grades involved, college professors, state curriculum supervisors, and researchers, was assembled to review the sets of questions identified for each percentile band. The committee was divided into three groups, one for each grade. Each group examined and analyzed questions that anchored at the 25th, 50th, and 90th percentiles to determine the specific U.S. history knowledge and abilities associated with each question.

Committee members were aiso provided with the sets of questions at each grade that "did not anchor" to inform their decisions about what students could do by seeing examples of what they could not do. Drawing on their knowledge of U.S. history, committee members were asked to summarize student performance by describing the knowledge, skills, and abilities demonstrated by students in each of the score bands.

The performance descriptions are cumulative; that is, the abilities described for the lower performing students are considered to be among the abilities of students performing at higher points on the scale. Therefore, the full description of students' U.S. history knowledge and abilities in the middle scale band would include those abilities described at the lower band. Similarly, the abilities of students performing at the higher scale band include the U.S. history abilities described for students at the middle and lower bands.

Profiling Students' Study Habits

Using the scale bands defined for the anchoring process described above, the profiling of students' study habits was accomplished by examining the responses of students within those bands to selected background questions. A complete presentation of students' responses to the three background variables highlighted in Chapter 5 are presented in Tables B.1 to B.3. The percentages that appear in the tables are conditional on the anchor scale point. That is, they are percentages of students who scored within a five percentile point range on either side of the specified scale point.

TABLE B.1

Responses of Students Near Selected Percentile Points to General Study Habit Questions Grade 4

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	25th Percentile Scale Range 171-187	50th Percentile Scale Range 205-215	90th Percentile Scale Range 246-263
Time Spent Each Day on Homework			
More Than 1 Hour	18 (2.2)	16 (2.9)	12 (2.2)
One Hour	25 (3.0)	31 (3.6)	34 (2.7)
One Half Hour or Less	41 (3.9)	38 (2.6)	37 (2.6)
Assigned/Don't Usually Do	4 (1.5)	2 (1.0)	2 (1.0)
Not Usually Assigned	12 (2.3)	13 (3.3)	15 (3.2)
Discuss Studies at Home			
Daily/Almost Daily	50 (3.8)	53 (5.2)	59 (3.0)
Once/Twice a Week	21 (2.4)	23 (3.8)	24 (2.2)
Once/Twice a Month	8 (2.0)	6 (1.4)	6 (1.6)
Never/Hardly Ever	21 (2.8)	18 (4.1)	11 (2.2)
Pages Read Each Day in School and for Homework			
More Than 20	19 (2.7)	22 (2.9)	26 (2.8)
16 to 20	14 (2.2)	17 (3.7)	18 (2.2)
11 to 15	13 (1.9)	15 (3.2)	16 (2.9)
6 to 10	24 (2.9)	23 (3.0)	25 (2.8)
5 or Fewer	30 (3.4)	23 (3.0)	15 (2.3)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

TABLE B.2

Responses of Students Near Selected Percentile Points to General Study Habit Questions Grade 8

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	25th Percentile Scale Range 233-244	50th Percentile Scale Range 257-265	90th Percentile Scale Range 292-308
Time Spent Each Day on Homework			
More Than 1 Hour	22 (1.8)	27 (2.2)	33 (3.3)
One Hour	35 (2.8)	38 (3.4)	38 (2.0)
One Half Hour or Less	23 (2.1)	20 (2.6)	21 (2.4)
Assigned/Don't Usually Do	11 (1.6)	8 (1.1)	4 (1.0)
Not Usually Assigned	8 (1.5)	7 (1.9)	4 (1.0)
Discuss Studies at Home			
Daily/Almost Daily	32 (2.7)	36 (2.1)	47 (2.4)
Once/Twice a Week	29 (2.1)	30 (2.0)	29 (2.2)
Once/Twice a Month	12 (1.7)	11 (1.8)	11 (1.2)
Never/Hardly Ever	27 (3.0)	23 (2.0)	14 (1.8)
Pages Read Each Day for School and Homework			
More Than 20	9 (1.6)	11 (2.0)	13 (1.8)
16 to 20	7 (1.0)	10 (1.6)	13 (1.8)
11 to 15	13 (1.5)	15 (2.1)	19 (2.1)
6 to 10	27 (2.1)	29 (2.7)	31 (1.9)
5 or Fewer	44 (3.0)	35 (2.5)	25 (2.6)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

TABLE B.3

Responses of Students Near Selected Percentile Points to General Study Habit Questions Grade 12

THE NATION'S REPORT CARD 1994 U.S. History Assessment

	25th Percentile Scale Range 259-270	50th Percentile Scale Range 284-292	90th Percentile Scale Range 319-335
Time Spent Each Day on Homework			
More Than 1 Hour	20 (2.3)	25 (2.7)	37 (2.4)
One Hour	30 (2.6)	31 (2.1)	27 (2.9)
One Half Hour or Less	23 (2.1)	24 (2.4)	24 (2.5)
Assigned/Don't Usually Do	9 (1.1)	7 (1.3)	7 (1.5)
Not Usually Assigned	19 (2.8)	13 (2.1)	5 (1.1)
Discuss Studies at Home			
Daily/Almost Daily	27 (2.7)	32 (3.7)	38 (2.6)
Once/Twice a Week	28 (2.3)	31 (2.3)	35 (2.0)
Once/Twice a Month	13 (1.9)	14 (1.8)	12 (2.1)
Never/Hardly Ever	32 (2.8)	24 (2.2)	15 (2.3)
Pages Read Each Day for School and Homework			
More Than 20	10 (1.5)	13 (1.6)	29 (2.1)
16 to 20	9 (1.9)	12 (1.7)	12 (1.9)
11 to 15	11 (1.9)	15 (2.2)	16 (1.8)
6 to 10	25 (1.8)	26 (3.0)	21 (2.2)
5 or Fewer	46 (2.2)	35 (3.0)	22 (2.7)

Differences between the groups may be partially explained by other factors not included in this table.

The standard errors of the estimated percentages appear in parentheses. It can be said with 95-percent certainty that, for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Percentages of students in the subgroups may not total 100 due to rounding.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 U.S. History Assessment.

Item Mapping Procedures

In order to map questions to particular points on the NAEP U.S. history subscales, a response probability convention had to be adopted that would divide those who had a higher probability of success from those who had a lower probability. Establishing a response probability convention has an impact on the mapping of assessment questions onto the U.S. history subscales. A lower boundary convention maps the U.S. history questions at lower points along the U.S. history scales, and a higher boundary convention maps the same questions at higher points along the scales. The underlying distribution of U.S. history skills in the population does not change, but the choice of a response probability convention does have an impact on the proportion of the student population that is reported as "able to do" the questions on the U.S. history scales.

There is no obvious choice of a point along the probability scale that is clearly superior to any other point. If the convention were set with a boundary at 50 percent, those above the boundary would be more likely to get a question right than get it wrong, while those below that boundary would be more likely to get the question wrong than right. While this convention has some intuitive appeal, it was rejected on the grounds that having a 50/50 chance of getting the question right shows an insufficient degree of mastery. If the convention were set with a boundary at 80 percent, students above the criterion would have a high probability of success with a question. However, many of the students below this criterion show some level of U.S. history ability that would be ignored by such a stringent criterion. In particular, those in the range between 50 and 80 percent correct would be more likely to get the question right than wrong, yet would not be in the group described as "able to do" the question.

In a compromise between the 50 percent and the 80 percent conventions, NAEP has adopted two related response probability conventions: (1) 74 percent for multiple-choice questions (to correct for the possibility of answering correctly by guessing), and (2) 65 percent for constructed-response questions (where guessing is not a factor). These probability conventions were established, in part, based on an intuitive judgment that they would provide the best picture of students' U.S. history knowledge and skills.

Some additional support for the dual conventions adopted by NAEP was provided by Huynh (1994).[1] He examined the IRT information provided by questions, according to the IRT model used in scaling NAEP questions. ("Information" is used here in a technical sense. See the *NAEP 1994 Technical Report* for details.) Following Bock (1972),[2] Huynh decomposed the item information into that provided by a correct response [P (θ) *I (θ)] and that provided by an incorrect response [(1-P (θ)) *I (θ)]. Huynh showed that the item information provided by a correct response to a constructed-response question is maximized at the point along the U.S. history scale at which two-thirds of the students get the question correct (for multiple-choice questions, information is maximized at the point at which 74 percent get the question correct). It should be noted, however, that maximizing the *item* information I (θ), rather than the information provided by a *correct response* [P (θ) *I (θ)], would imply an item mapping criterion closer to 50 percent.

Endnotes

1. Huynh, H. (1994, October). *Some technical aspects of standard setting.* Paper presented at the Joint Conference on Standard Setting for Large-Scale Assessment, Washington, DC.
2. Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more latent categories. *Psychometrika, 37*, 29-51.

APPENDIX C

Sample Questions from the NAEP 1994 U.S. History Assessment

This appendix presents additional sample questions and student responses selected for each grade to exemplify the range of exercises included in the NAEP 1994 U.S. history assessment. (A different set of sample questions and student responses are presented in Chapter 1, and an entire sample block of questions can be found in the *NAEP 1994 U.S. History: A First Look* report, pages 29 to 54.) For each question, the historical theme and historical period being addressed is indicated. For multiple-choice questions, the correct answer is marked (►). For constructed-response questions, an abbreviated scoring rubic is provided. The sample student responses have been reproduced from assessment booklets and represent typical student performance.

The table accompanying each sample question presents two types of percentages: (1) the overall percentage of students within a grade who answered the question successfully, and (2) the percentages of students within each of the achievement level intervals — *Basic, Proficient,* and *Advanced* — who answered the question successfully. The percentages for students within the *Advanced* achievement level interval are not presented, however, because of the small sample sizes. The percentages of students *below Basic* who successfully answered the questions are also included in the tables. (Sample size criteria for reporting results are described in Appendix A.)

What is the main reason the Pilgrims and Puritans came to America?

- ▶ A To practice their religion freely
- B To make more money and live a better life
- C To build a democratic government
- D To expand the lands controlled by the king of England

Historical Theme:
The Gathering and Interactions of Peoples, Cultures, and Ideas

Historical Period:
Colonization, Settlement, and Communities (1607 to 1763)

Grade 4	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	*Below Basic* 194 and below*	*Basic* 195-242*	*Proficient* 243-275*	*Advanced* 276 and above*
41 (1.3)	20 (2.0)	44 (2.8)	73 (4.9)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Your teacher has asked you to teach your classmates about ONE of these famous places where an important event in American history happened:

the Alamo

Pearl Harbor

Gettysburg

Roanoke Island

My famous place in American history is_____

Write down three facts about the place that you have chosen that will help you teach your classmates about that place.

Historical Theme:

The Changing Role of America in the World

Historical Periods:

Three Worlds and Their Meeting in the Americas (Beginnings to 1607)

Expansion and Reform (1801 to 1861)

Modern America and the World Wars (1914 to 1945)

Responses to this question were scored according to a four-level rubric as 1) Inappropriate, 2) Partial, 3) Essential, and 4) Complete.

Grade 4	Percentage "Essential" or "Complete" within Achievement Level Intervals			
Overall Percentage Essential or Complete	*Below Basic* 194 and below*	*Basic* 195-242*	*Proficient* 243-275*	*Advanced* 276 and above*
19 (1.4)	2 (0.8)	17 (2.4)	57 (4.2)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (Score of 3):

My famous place in American history is Gettysburg

Write down three facts about the place that you have chosen that will help you teach your classmates about that place.

Fact 1 It was during the civil war.

Fact 2 Many people died in this battle.

Fact 3 It ended the civil war.

An **Essential** response (score of 3) gives two facts that are relevant to the particular place and that would help another person understand the place.

Sample Response (Score of 4):

My famous place in American history is Pearl Harbor

Write down three facts about the place that you have chosen that will help you teach your classmates about that place.

Fact 1 The Japenes groped a bome on Pearl Harbor

Fact 2 The bome distroed lots of ships

Fact 3 The U.S.A. fote back

A **Complete** response (score of 4) gives three facts that are relevant to the particular place and that would help another person understand the place, such as that the bombing of Pearl Harbor caused the U.S. to enter WWII, or that the battle of Gettysburg was a turning point in the Civil War.

Which of these was one of the thirteen colonies that fought the American Revolution against the British?

A Illinois

B California

▶ C New York

D Texas

Historical Theme:

The Changing Role of America in the World

Historical Period:

The Revolution and the New Nation (1763 to 1815)

Grade 4 Overall Percentage Correct	Percentage Correct within Achievement Level Intervals			
	Below Basic 194 and below*	*Basic* 195-242*	*Proficient* 243-275*	*Advanced* 276 and above*
32 (1.5)	22 (2.1)	28 (2.3)	59 (4.7)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

What is the purpose of the Bill of Rights?

A To say how much Americans should pay in taxes

▶ B To protect freedoms like freedom of speech

C To describe the jobs of the President and Congress

D To make Washington, D.C., the capital of the United States

Historical Theme:

Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

Historical Period:

The Revolution and the New Nation (1763 to 1815)

Grade 4 Overall Percentage Correct	Percentage "Essential" or Complete within Achievement Level Intervals			
	Below Basic 194 and below*	*Basic* 195-242*	*Proficient* 243-275*	*Advanced* 276 and above*
45 (1.4)	32 (2.2)	46 (1.9)	65 (4.8)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Name two kinds of work women do today that they could not do 100 years ago.

1. _____
   _____
2. _____
   _____

Explain why the kinds of work women do today are different from the kinds of work women did 100 years ago.

_____
_____
_____

Historical Theme:

The Gathering and Interactions of Peoples, Cultures, and Ideas

Historical Period:

The Development of Modern America (1865 to 1920)

Responses to this question were scored according to a three-level rubric as 1) Inappropriate, 2) Partial, and 3) Appropriate.

Grade 4 Overall Percentage Appropriate	Percentage "Appropriate" within Achievement Level Intervals			
	Below Basic 194 and below*	*Basic* 195-242*	*Proficient* 243-275*	*Advanced* 276 and above*
11 (0.9)	1 (0.6)	11 (1.1)	30 (3.7)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (Score of 3)

Name two kinds of work women do today that they could not do 100 years ago.

1. They can be a doctor and they couldn't 100 years ago.

2. They can be a lawyer and they couldn't 100 years ago.

Explain why the kinds of work women do today are different from the kinds of work women did 100 years ago.

The kinds of work women do today are different than the kinds of work women did 100 years ago because 100 years ago people thought some jobs were men's jobs and that ladys couldn't do them.

An **Appropriate** response (score of 3) identifies two types of work, and shows some understanding of the fact that women have more opportunities today.

Between 1960 and 1990, what invention most changed the way people in the United States worked?

A The typewriter

▶B The computer

C The superconductor

D The radio

Historical Theme:

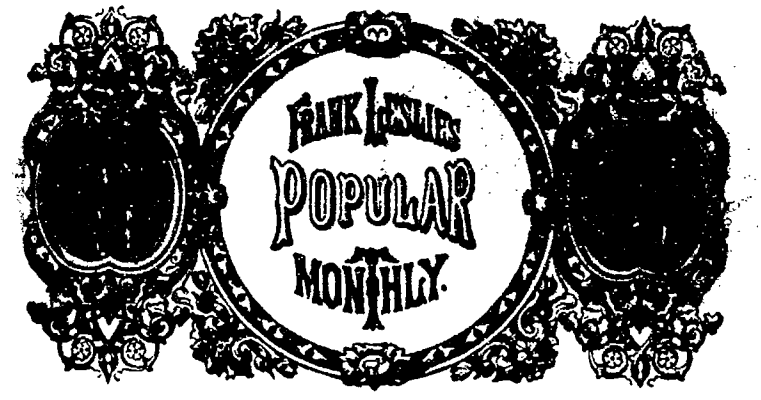
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

Historical Period:

Contemporary America (1945 to Present)

Grade 8	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	***Below Basic*** **251 and below***	***Basic*** **252-293***	***Proficient*** **294-326***	***Advanced*** **327 and above***
78 (1.2)	**65 (2.6)**	**83 (1.6)**	**94 (2.2)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.



Vol. II.—No 3 SEPTEMBER, 1876 $2 50

WHAT SHALL WE DO WITH OUR INDIANS?

Philadelphia Museum of Art:
Purchased: Lola Downin Peck Fund.

Look at the magazine cover. What historical events would have led this question and picture to appear on the cover of a popular magazine in 1876?

What attitudes displayed toward American Indians by other Americans are suggested by this magazine cover?

Historical Theme:
The Gathering and Interactions of Peoples, Cultures, and Ideas

Historical Period:
The Development of Modern America (1865 to 1920)

Responses to this question were scored according to a four-level rubric as 1) Inappropriate, 2) Partial, 3) Essential, and 4) Complete.

Grade 8	Percentage "Essential" or "Complete" within Achievement Level Intervals			
Overall Percentage Essential or Complete	*Below Basic* 251 and below*	*Basic* 252-293*	*Proficient* 294-326*	*Advanced* 327 and above*
25 (1.2)	6 (1.4)	32 (2.0)	59 (4.4)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (Score of 3)

Look at the magazine cover. What historical events would have led this question and picture to appear on the cover of a popular magazine in 1896?

The genocide of Indians while settlers enjoyed their new land

What attitudes displayed toward American Indians by other Americans are suggested by this magazine cover?

That some people felt sorry for the Native Americans

An **Essential** response (score of 3) answers both parts of the question, although one part gives a general statement without providing any specifics. For instance the response may state that the issue is addressed on a magazine cover because "there were lots of arguments between the government and the Indians then"; or, the response may describe the attitude of many other Americans by saying "they did not like the Sioux Indians."

Sample Response (Score of 4)

Look at the magazine cover. What historical events would have led this question and picture to appear on the cover of a popular magazine in 1896?

Indian uprisings and ambushes of frontier farmers army putting Indians on reservation

What attitudes displayed toward American Indians by other Americans are suggested by this magazine cover?

that they were less than human and far behind rest of the world and they were a nuisance to frontier farmers and settlers and had to be put away on reservations

A **Complete** response (score of 4) answers both parts of the question and provides specifics. It may, for example, discuss disputes about western lands, and the issue of reservations vs. assimilation.

We hold these truths to be self-evident: That all men are created equal; that they are endowed by their Creator with certain unalienable rights; that among these are life, liberty, and the pursuit of happiness. That, to secure these rights, governments are instituted among men, deriving their just powers from the consent of the governed; that, whenever any form of government becomes destructive of these ends, it is the right of the people to alter or to abolish it, and to institute a new government.

—— 1776

The primary author of the document was

A George Washington

B John Marshall

C Robert E. Lee

►D Thomas Jefferson

Historical Theme:

Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

Historical Period:

The Revolution and the New Nation (1763 to 1815)

Grade 8 Overall Percentage Correct	Percentage Correct within Achievement Level Intervals			
	Below Basic 251 and below*	*Basic* 252-293*	*Proficient* 294-326*	*Advanced* 327 and above*
71 (1.6)	55 (2.4)	79 (2.2)	93 (2.0)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

During the Constitutional Convention, there was an important debate between large and small states about representation. What were the main issues in this debate?

_____
_____
_____

How did the Connecticut (or Great) Compromise resolve this debate?

_____
_____
_____

Historical Theme:

Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies

Historical Period:

The Revolution and the New Nation (1763 to 1815)

Responses to this question were scored according to a four-level rubric as 1) Inappropriate, 2) Partial, 3) Essential, and 4) Complete.

Grade 8	**Percentage "Essential" or "Complete" within Achievement Level Intervals**			
Overall Percentage Essential or Complete	***Below Basic*** **251 and below***	***Basic*** **252-293***	***Proficient*** **294-326***	***Advanced*** **327 and above***
10 (1.1)	**0 (0.3)**	**6 (1.6)**	**43 (4.7)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (Score of 3)

During the Constitutional Convention, there was an important debate between large and small states about representation. What were the main issues in this debate?

How to divide the houses bigger states wanted by population the smaller states wanted one per state.

How did the Connecticut (or Great) Compromise resolve this debate?

They came up with 2 houses.

An **Essential** response (score of 3) identifies the debate about representation, but does not fully explain the mechanism of the resolution.

Sample Response (Score of 4)

During the Constitutional Convention, there was an important debate between large and small states about representation. What were the main issues in this debate?

The larger states wanted more representation because they thought since they were bigger they should have more control. The smaller states did not agree. They wanted equal representation. This disagreement led to The Great Compromise

How did the Connecticut (or Great) Compromise resolve this debate?

The Great Compromise settled this debate by establishing two house, the Senate and House of Representatives. The Senate would have 2 representatives no matter how large to state and the House would have a representatives according to state population by the three-fifths Amendment.

A **Complete** response (score of 4) fully explains the debate about representation, and the solution of the bicameral legislature.

The Lend-Lease Act, the Yalta Conference, and the dropping of the atomic bomb on Hiroshima are all associated with the

A First World War

▶B Second World War

C Korean War

D Vietnam War

Historical Theme:
The Changing Role of America in the World

Historical Period:
Modern America and the World Wars (1914 to 1945)

Grade 8 Overall Percentage Correct	Percentage Correct within Achievement Level Intervals			
	Below Basic 251 and below*	*Basic* 252-293*	*Proficient* 294-326*	*Advanced* 327 and above*
41 (1.3)	29 (2.2)	43 (2.4)	69 (3.5)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

O Freedom!
O Freedom!
O Freedom over me!
And before I'd be a slave,
I'd be buried in my grave,
And go home to my Lord and be free!

The song was associated with

A the temperance movement

▶B the civil rights movement

C pioneers on the Oregon Trail

D farmers in the Dust Bowl during the Great Depression

Historical Theme:
The Gathering and Interactions of Peoples, Cultures, and Ideas

Historical Period:
Contemporary America (1945 to present)

Grade 8	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	Below Basic 251 and below*	Basic 252-293*	Proficient 294-326*	Advanced 327 and above*
80 (1.1)	67 (2.2)	86 (1.8)	93 (2.6)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

The Great Awakening of the 1730's was important because it led people in the American colonies to

A increase toleration for Roman Catholics

B examine the different positions of men and women in society

C reaffirm that God gave kings their right to rule

➤D question the authority of church and government leaders

Historical Theme:
The Gathering and Interactions of Peoples, Cultures, and Ideas

Historical Period:
Colonization, Settlement, and Communities (1607 to 1763)

Grade 12	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	***Below Basic*** **293 and below***	***Basic*** **294-324***	***Proficient*** **325-354***	***Advanced*** **355 and above***
56 (1.5)	**42 (2.2)**	**74 (2.5)**	**80 (4.4)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

"Our reconstruction measures were radically defective because they failed to give the ex-slaves any land."

—— Frederick Douglass

Describe briefly the way in which Douglass's statement helps explain the rise of sharecropping in the South after the Civil War. In your answer, be sure to define the term sharecropping.

Historical Theme:

The Gathering and Interactions of Peoples, Cultures, and Ideas

Historical Period:

Crisis of the Union: Civil War and Reconstruction (1850 to 1877)

Responses to this question were scored according to a four-level rubric as 1) Inappropriate, 2) Partial, 3) Essential, and 4) Complete.

Grade 12	Percentage "Essential" or "Complete" within Achievement Level Intervals			
Overall Percentage Essential or Complete	***Below Basic*** **293 and below***	***Basic*** **294-324**	***Proficient*** **325-354***	***Advanced*** **355 and above***
18 (1.4)	**5 (0.8)**	**27 (3.0)**	**57 (5.9)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (score of 3)

Describe briefly the way in which Douglass's statements helps explain the rise of sharecropping in the South after the Civil War. In your answer, be sure to define the term sharecropping.

The rise of sharecropping was increased after the Civil War. There are many reasons for that. First of all sharecropping is using a persons land while someone else takes care of it. They farm it and make sure the crops are doing alright it was usually the black men who sharecropped with the white men. This made it seem as though the blacks were still slaves because they couldn't buy their own land and the owners were still mean and wouldn't give them as much money as they should have.

An **Essential** response (score of 3) correctly defines sharecropping, and links it to difficult economic conditions faced by slaves.

Sample Response (score of 4)

Describe briefly the way in which Douglass's statements helps explain the rise of sharecropping in the South after the Civil War. In your answer, be sure to define the term sharecropping.

Sharecropping, tenanting out land for a set portion of the crop, rose up after the Civil War because ex-slaves had no land. Ex-slaves had to find work, but no work was to be had since they could not farm their own land. If the ex-slaves had had land, small farms would have been run preventing the problems associated with sharecropping such as people being tied to the land by debt. Instead of this, landowners simply parceled off the large tracts of land they could not farm alone creating a sharecropping based economy.

A **Complete** response (score of 4) correctly defines sharecropping. and links it to Reconstruction policies, or links the failure to give slaves land to a cycle of poverty that led to sharecropping.

Document G - ***Excerpt from Franklin D. Roosevelt's First Inaugural Address (1933)***

> This great Nation will endure as it has endured, will revive and will prosper. So, first of all, let me assert my firm belief that the only thing we have to fear is fear itself — nameless, unreasoning, unjustified terror which paralyzes needed efforts to convert retreat into advance.
>
> A host of unemployed citizens face the grim problem of existence, and an equally great number toil with little return.
>
> Our greatest primary task is to put people to work. This is not an unsolvable problem if we face it wisely and courageously. It can be accomplished in part by direct recruiting by the Government itself, . . . accomplishing greatly needed projects to stimulate and reorganize the use of our natural resources.
>
> — Franklin D. Roosevelt's First Inaugural Address (1933)

Document I - ***Excerpt from Franklin D. Roosevelt's Second Inaugural Address (1937)***

> We have always known that heedless self-interest was bad morals; we know now that it is bad economics. Out of the collapse of a prosperity whose builders boasted their practicality has come the conviction that in the long run economic morality pays. We are beginning to wipe out the line that divides the practical from the ideal; and in so doing we are fashioning an instrument of unimagined power for the establishment of a morally better world.
>
> Let us ask again. Have we reached the goal of our vision of that fourth day of March, 1933? Have we found our happy valley?
>
> I see a great nation, upon a great continent, blessed with a great wealth of natural resources.
>
> But here is the challenge to our democracy. In this nation I see tens of millions of its citizens a substantial part of its whole population who at this very moment are denied the greater part of what the very lowest standards of today call the necessities of life.
>
> I see one-third of a nation ill-housed, ill-clad, ill-nourished.
>
> We are determined to make every American citizen the subject of his country's interest and concern; and we will never regard any faithful, law-abiding group within our borders as superfluous.
>
> — Franklin D. Roosevelt's Second Inaugural Address (1937)

Reread Roosevelt's 1933 and 1937 inaugural speeches. Would you characterize Roosevelt as more pessimistic or more optimistic in 1937 than in 1933? Explain why. Be sure to refer to the documents to support your answer

_____

_____

_____

Historical Theme:

Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

Historical Period:

Modern America and the World Wars (1914 to 1945)

Responses to this question were scored according to a three-level rubric as 1) Inappropriate, 2) Partial, and 3) Appropriate.

Grade 12	Percentage "Appropriate" within Achievement Level Intervals			
Overall Percentage Appropriate	*Below Basic* 293 and below*	*Basic* 294-324*	*Proficient* 325-354*	*Advanced* 355 and above*
22 (1.5)	9 (1.4)	31 (3.6)	57 (5.7)	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (score of 3)

Reread Roosevelt's 1933 and 1937 inaugural speeches. Would you characterize Roosevelt as more pessimistic or more optimistic in 1937 than in 1933? Explain why. Be sure to refer to the documents to support your answer.

It made pessimistic. In 1937 he seems to be criticizing the selfishness of his country and questioning its morality whereas in 1933 he was telling his country to fear not, and to hang tough so they could kill the depression,

An **Appropriate** response (score of 3) defends either point of view, with specific references to the documents. The student may, for example, cite either Roosevelt's description of accomplishments, or his questioning of America's moral fiber.

The Monroe Doctrine was intended to

A promote United States trade with China

B help keep the peace in Europe

►C discourage European involvement in the Americas

D protect United States business in Japan and Korea

Historical Theme:
The Changing Role of America in the World

Historical Period:
Expansion and Reform (1801 to 1861)

Grade 12	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	***Below Basic*** **293 and below***	***Basic*** **294-324**	***Proficient*** **325-354***	***Advanced*** **355 and above***
41 (1.5)	**27 (1.9)**	**54 (3.2)**	**80 (4.2)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

SOVIETS LAUNCH FIRST MAN-MADE SATELLITE INTO ORBIT

Frank Williams in the Detroit Free Press

What did the United States government do in response to the event referred to in the cartoon and headline?

A The government decided to seek peace immediately and to end the Cold War.

B The government banned civilian contact between United States and Soviet citizens.

▶C The government decided to spend more on both scientific education and the military.

D The government requested that the United Nations prohibit Soviet space exploration.

Historical Theme:
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

Historical Period:
Contemporary America (1945 to Present)

Grade 12	Percentage Correct within Achievement Level Intervals			
Overall Percentage Correct	***Below Basic*** **293 and below***	***Basic*** **294-324***	***Proficient*** **325-354***	***Advanced*** **355 and above***
74 (1.4)	**61 (2.2)**	**90 (2.2)**	**99 (0.9)**	***

*NAEP U.S. history composite scale range ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

AVERAGE FARM SIZE AND TOTAL NUMBER OF FARMS		
Year	Farm Size	Number of Farms
1900	150 acres	6,250,000
1980	425 acres	2,225,000

Summarize the changes shown in the table above.

_____

_____

_____

Explain how one invention or development helped cause the changes you have described.

_____

_____

_____

Historical Theme:

Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment

Historical Period:

The Development of Modern America (1865 to 1920)

Responses to this question were scored according to a three-level rubric as 1) Inappropriate, 2) Partial, and 3) Appropriate.

Grade 12	**Percentage "Appropriate" within Achievement Level Intervals**			
Overall Percentage Appropriate	***Below Basic*** **293 and below***	***Basic*** **294-324***	***Proficient*** **325-354***	***Advanced*** **355 and above***
29 (1.5)	**16 (1.5)**	**41 (3.3)**	**58 (4.1)**	***

*NAEP U.S. history composite scale range. ***Sample size insufficient to permit a reliable estimate (See Appendix A).
The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample.

Sample Response (Score of 3)

Summarize the changes shown in the table above.

Between the years 1900 & 1980, the size of farms increased, while the number of them decreased.

Explain how one invention or development helped cause the changes you have described.

One major development was the tractor. This enabled the farmers to get a lot more work done a lot quicker. They could have bigger farms & get the work finished with the tractor as fast as having a small farm and working without one.

An **Appropriate** response (score of 3) indicates that there are fewer, larger farms, and offers an acceptable reason for this. Students may, for example, refer to the growth of agribusinesses or changes in technology.

*U.S. Government Printing Office: 1996 - 410-618/50317

ACKNOWLEDGMENTS

The NAEP 1994 United States history assessment was a collaborative effort among staff from the National Center for Education Statistics (NCES), the National Assessment Governing Board (NAGB), Educational Testing Service (ETS), Westat, and National Computer Systems (NCS). The program benefited from the contributions of hundreds of individuals at the state and local levels — governors, chief state school officers, state and district test directors, state coordinators, and district administrators — who tirelessly provided their wisdom, experience, and hard work. Most importantly, NAEP is grateful to students and school staff who made the assessment possible.

The assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. The NCES staff — particularly Jeanne Griffith, Gary Phillips, Steve Gorman, Susan Ahmed, Peggy Carr, Sharif Shakrani, Larry Ogle, Sahar Akhtar, and Maureen Treacy — worked closely and collegially with ETS, Westat, and NCS staff and played a crucial role in all aspects of the program. The NAEP 1994 assessments and reports also benefitted from the consistent support and guidance of Emerson Elliott, past commissioner of NCES. The members of the National Assessment Governing Board (NAGB) and the NAGB staff provided invaluable advice and guidance throughout. NAEP also owes a debt of gratitude to the numerous panelists and consultants who provided their expertise and worked so conscientiously on developing the assessment.

The NAEP project at ETS is directed by Paul Williams and resides in the Center for the Assessment of Educational Progress (CAEP), managed by Archie Lapointe and Paul Williams. Steve Lazer managed test development activities, and Alexandra Beatty and Hilary Persky worked with the United States History Item Development committee to develop the assessment instruments. Jules Goodison managed the operational aspects together with John Olson, and sampling and data collection activities were carried out by Westat under the direction of Rene Slobasky, Nancy Caldwell, and Keith Rust. Printing, distribution, scoring, and processing activities were conducted by NCS, under the supervision of Judy Moyer, Brad Thayer, Mathilde Kennel, Linda Reynolds, and Barbara Price.

Statistical and psychometric activities for the assessments were led by Eiji Muraki under the direction of Eugene Johnson, John Mazzeo, and Jim Carlson. Bruce Kaplan and Jim Ferris performed the United States history analyses. Karen Miller and John Mazzeo contributed substantially to report design activities. Mary Michaels and Sharon Davis-Johnson oversaw the production aspects, and Loretta Casalaina, James Rura, Roderick Rudder, and Sheri Barnes provided further design and composition assistance.

Many thanks are due to the numerous reviewers, both internal and external to ETS and NCES. The comments and critical feedback provided by the following reviewers are reflected in the final version of this report: John Patrick, Phoebe Speck, Mary Lyn Bourque, Larry Feinberg, Mary Frase, Gary Phillips, Sharif Shakrani, Aurora D'Amico, Shi-Chang Wu, Andrew Kolstad, Susan Ahmed, Bruce Kaplan, Karen Miller, Nada Ballator, Steve Lazer, and John Mazzeo. Beverly Cisney and Alice Kass provided the excellent desktop publishing skills essential to the project.

UNITED STATES
DEPARTMENT OF EDUCATION
WASHINGTON, DC 20208-5653

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

Fourth Class Special
Special Handling

NCES 96–085