Appendix A
Overview of Procedures Used for the NAEP 2002 Writing Assessment

This appendix provides an overview of the NAEP 2002 writing assessment's primary components—framework, development, administration, scoring, and analysis. A more extensive review of the procedures and methods used in the writing assessment will be included in the NAEP 2002 technical documentation section of the NAEP web site (*http://nces.ed.gov/nationsreportcard*).

The NAEP Writing Assessment

The NAEP 2002 writing assessment is based on the 1998 writing assessment framework.[1] The framework's purpose is to provide, based on the expert opinions of writing educators and researchers, a definition of writing upon which the NAEP writing assessment can be based. The framework development process 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. The process was managed by the Center for Research on Evaluation, Standards, and Student Testing (CRESST) for the National Assessment Governing Board (NAGB), and the exercise specifications were developed under contract by American College Testing (ACT) for NAGB.

[1] National Assessment Governing Board. (2002). *Writing Framework for the 1998 National Assessment of Educational Progress*. Washington, DC: Author.

The writing framework delineates six major objectives to organize the design of the assessment.

- Students should write for three major purposes: narrative, informative, and persuasive. While other types of writing could have been included, the developers of the framework believed that, for the purpose of monitoring student achievement (as opposed to creating individual diagnostic assessments), three broad types of writing were appropriate.
- Students should be able to write on a variety of tasks (letters, essays, stories, reports) and for different audiences (peers, school or government officials, business representatives).
- Student writing should be prompted by a variety of stimulus materials (letters, poems, graphics, reports) under varying time constraints.
- Because writing is a dynamic process through which the writer constructs meaning, students should develop their own writing processes, including methods for drafting, evaluating, revising, and editing ideas and forms of expression. Students are to be given planning space in the test materials to generate ideas for drafts. In addition, they are given a pamphlet with suggestions for planning, revising, and editing. All NAEP student responses, given assessment time constraints, are to be evaluated as first drafts.
- Students should display effective choices in the organization of their writing. Further, they should be able to illustrate and elaborate their ideas and should use appropriate conventions of English. All of these characteristics are to be part of the evaluation of student writing.
- Students should value writing as a communicative activity.

Figure A.1 gives examples of various writing tasks similar to those included in the assessment at grades 4, 8, and 12. Included in the figure are descriptions of sample tasks that illustrate how each purpose for writing is assessed.

Figure A.1 Illustrative examples of writing tasks, by purpose for writing, grades 4, 8, and 12

Purpose for writing	Grade 4	Grade 8	Grade 12
Narrative	Provide visual stimuli of a season of the year. Ask students to write a letter to a grandparent telling the story of an interesting personal experience related to the season.	Provide visual stimuli. Ask students to write an article for a sports magazine telling the story of a time when they participated in a hobby or skill they enjoyed.	Provide an appropriate quotation. Ask students to write a letter to a friend telling the story of a time in their lives when they had to make an important decision.
Informative	Provide an appropriate quotation. Ask students to explain in an essay to their English teacher how a person (parent, teacher, friend) has influenced them in an important way.	Provide a series of brief journal entries from another historical time. Ask students to explain what is revealed about the person who wrote the entries.	Provide quotations from a political campaign. Ask students to choose one and in an essay inform their social studies teacher what it means in the context of the campaign.
Persuasive	Provide visual stimuli of an animal. Ask students to convince their parents/ guardians of an animal that would make the best pet.	Provide brief reviews, as models, of a film, TV program, or book. Ask students to write a review for the school newspaper that will convince other students to watch a favorite film or TV program or read a favorite book.	Provide a quotation on education in the United States. Ask students to write a letter to the editor of their local newspaper taking a position on some aspect of education and support it from their own experiences.

SOURCE: National Assessment Governing Board. *Writing Framework and Specifications for the 1998 National Assessment of Educational Progress*. Washington, DC: Author.

In addition to the six objectives, the framework specifies the percentage of the writing tasks in the assessment that should be devoted to each of the three writing purposes—narrative, informative, and persuasive. The actual percentage distributions of writing tasks in the assessment are listed in table 1.1 of chapter 1. The table shows the number of tasks at each grade level for each purpose. Each task received equal weight in the composition of the NAEP scale for each grade. These target percentages vary by grade level according to what is deemed developmentally appropriate for each grade, as stated in the writing framework.

The Assessment Design

Each student who participated in the writing assessment received a booklet containing two 25-minute writing tasks. All student responses to the writing tasks were rated according to a six-level scoring guide. In addition, the test booklets contained general background questions and writing-specific background questions.

The assessment design allowed for maximum coverage of the writing domain at each grade, while minimizing the time burden for any one student. This was accomplished through the use of matrix sampling of tasks, in which each student was given only 2 of the 20 tasks at each grade level. Representative samples of students responded to each task, so that the aggregate results across the entire assessment allow broad reporting of writing abilities for the targeted population.

In addition to matrix sampling, the assessment design utilized a procedure for distributing blocks across booklets that controlled for position and context effects. Students received different blocks of tasks in their booklets according to a procedure called "partially balanced incomplete block (PBIB) spiraling." The procedure assigned blocks of questions in a manner that balanced the positioning of blocks across booklets and balanced the pairing of blocks within booklets according to purposes for writing. Blocks were balanced within each purpose for writing and were partially balanced across purposes for writing. (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 booklets.)

In addition to the student assessment booklets, three other instruments provided data relating to the assessment—a teacher questionnaire, a school questionnaire, and a students with disabilities/limited English proficient student (SD/LEP) questionnaire. The SD/LEP questionnaire was completed by a school staff member knowledgeable about those students who were selected to participate in the assessment and who were identified as having an Individualized Education Program (IEP) or equivalent plan or being limited English proficient (LEP). An SD/LEP questionnaire was completed for each identified student regardless of whether the student participated in the assessment. Each SD/LEP questionnaire asked about the student and the special programs in which he or she participated.

NAEP Samples

National Sample

The national results presented in this report are based on nationally representative probability samples of fourth-, eighth-, and twelfth-grade students. At grades 4 and 8, the national sample consisted of the combined sample of students assessed in each participating state, plus an additional sample from the states that did not participate in the state assessment, as well as a private school sample. This represents a change from the 1998 assessment in which the national and state samples were independent. At grade 12, the sample was chosen using a stratified two-stage design that involved sampling students from selected schools (public and nonpublic) across the country.

Each selected school that participated in the assessment and each student assessed represents a portion of the population of interest. Sampling weights are needed to make valid inferences between the student samples and the respective populations from which they were drawn. Sampling weights account for disproportionate representation of students from different states, and students who attend nonpublic schools. Sampling weights also account for lower sampling rates for very small schools and are used to adjust for school and student nonresponse.[2]

Table A.1 provides a summary of the 2002 national school and student participation rates for the writing assessment sample. Participation rates are presented for public and nonpublic schools both individually and combined. Four different rates are presented; the first rate is a student-centered weighted percentage of schools participating in the assessment before substitution of demographically similar schools.[3] This rate is based only on the schools that were initially selected for the assessment. The numerator of this rate is the estimated number of students represented by the initially selected schools that participated in the assessment. The denominator is the estimated number of students represented by the initially selected schools that had eligible students enrolled.

The second school participation rate is a student-centered weighted participation rate after substitution. The numerator of this rate is the estimated number of students represented by the participating schools, whether originally selected or selected as a substitute for a school that chose not to participate. The denominator is the estimated number of students represented by the initially selected schools that had eligible students enrolled (this is the same as that for the weighted participation rate for the sample of schools before substitution). Because of the common denominators, the weighted participation rate after substitution is at least as great as the weighted participation rate before substitution.

The third school participation rate is a school-centered weighted percentage of schools participating in the assessment, before substitution of demographically similar schools. This rate is based only on the schools that were initially selected for the assessment. The numerator of this rate is the estimated number of schools represented by the initially selected schools that participated in the assessment. The denominator is the estimated number of schools represented by the initially selected schools that had eligible students enrolled.

The fourth school participation rate is a school-centered weighted participation rate after substitution. The numerator is the estimated number of schools represented by the participating schools, whether originally selected or selected as a substitute for a school that did not participate. The denominator is the estimated number of schools, represented by the initially selected schools that had eligible students enrolled.

[2] Additional details regarding the design and structure of the national and state samples will be included in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

[3] The initial base sampling weights were used in weighting the percentages of participating schools and students. An attempt was made to preselect one substitute school for each sampled public school and one for each sampled Catholic school, and one for each sampled nonpublic school other than Catholic. To minimize bias, a substitute school resembled the original selection as much as possible in affiliation, type of location, estimated number of grade-eligible students, and minority composition.

The student-centered and school-centered school participation rates differ if school participation is associated with the size of the school. If the student-centered rate is higher than the school-centered rate, this indicates that larger schools participated at a higher rate than smaller schools. The converse applies also.

Also presented in table A.1 are weighted student participation rates. The numerator of this rate is the estimated number of students who are represented by the students assessed (in either an initial session or a makeup session). The denominator of this rate is the estimated number of students represented by the eligible sampled students in participating schools.

For the grade 12 national sample, where school and student response rates did not meet NCES standards, an extensive analysis was conducted that examined, among other factors, the potential for nonresponse bias at both the school and student level. Results of these analyses, as well as nonresponse bias analyses for the grades 4 and 8 national samples, will be included in the technical documentation.

Table A.1 National school and student participation rates, by type of school, grades 4, 8, and 12: 2002

	School participation					Student participation	
	Student weighted		School weighted				
	Percentage before substitution	Percentage after substitution	Percentage before substitution	Percentage after substitution	Number of schools participating after substitution	Student weighted percentage	Number of students assessed
Grade 4							
Combined national	84	85	80	83	5,518	94	139,198
Public	85	85	84	85	5,067	94	132,753
Nonpublic	74	81	69	77	451	95	5,383
Grade 8							
Combined national	82	83	74	78	4,706	92	118,516
Public	83	84	80	81	4,208	91	112,485
Nonpublic	68	76	65	74	498	95	5,499
Grade 12							
Combined national	74	75	68	71	725	74	18,532
Public	76	76	73	74	443	72	14,291
Nonpublic	55	59	53	60	282	88	4,241

NOTE: The number of students in the combined national total at grades 4 and 8 includes students in the Department of Defense domestic schools located within the U.S. and Bureau of Indian Affairs schools that are not included as part of either the public or nonpublic totals.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

State Samples

The results provided in this report of the 2002 state assessment in writing are based on state-level samples of fourth- and eighth-grade public-school students. The samples were selected using a two-stage sample design that first selected schools within participating states and other jurisdictions and then students within schools. The samples were weighted to allow valid inferences about the populations of interest. Participation rates for jurisdictions were calculated the same way that rates were computed for the nation. Tables A.2 and A.3 contain the number of participating schools and students, as well as weighted school and student participation rates for the state samples at grades 4 and 8 respectively.

Table A.2 School and student participation rates, grade 4 public schools: By state, 2002

Grade 4	School participation			Student participation		Overall participation rate	
	Student weighted						
	Percentage before substitution	Percentage after substitution	Number of schools participating after substitution	Student weighted percentage	Number of students assessed	Before substitution	After substitution
Nation (Public)	85	85	5,067	94	132,753	80	80
Alabama	84	96	108	95	3,575	80	92
Arizona	91	91	105	91	3,073	83	83
Arkansas	99	99	107	94	2,779	93	93
California ‡	72	72	143	95	3,979	68	68
Connecticut	100	100	108	95	3,174	95	95
Delaware	100	100	86	94	3,950	94	94
Florida	100	100	103	95	3,210	95	95
Georgia	100	100	152	95	4,852	95	95
Hawaii	100	100	111	96	3,602	96	96
Idaho	87	87	98	95	2,722	82	82
Illinois ‡	57	57	117	93	3,053	53	53
Indiana	99	99	112	94	3,398	93	93
Iowa ‡	77	77	86	95	1,948	73	73
Kansas ‡	73	73	84	96	1,900	70	70
Kentucky	96	96	106	96	3,219	92	92
Louisiana	99	99	116	96	3,270	95	95
Maine	88	88	98	94	1,937	83	83
Maryland	100	100	105	93	2,791	93	93
Massachusetts	100	100	111	95	3,141	95	95
Michigan	98	99	110	92	2,970	90	91
Minnesota ‡	77	77	84	95	2,574	73	74
Mississippi	95	95	104	95	2,985	90	90
Missouri	94	100	113	94	2,963	89	94
Montana ‡	75	75	77	95	1,332	71	71
Nebraska	95	95	87	96	1,497	91	91
Nevada	100	100	113	93	3,474	93	93
New Mexico	93	93	104	94	2,348	87	87
New York ‡	77	77	90	91	2,370	70	70
North Carolina	100	100	113	94	3,366	94	94
North Dakota ‡	82	82	158	96	2,368	79	79
Ohio	95	95	107	93	2,688	89	89
Oklahoma	99	99	132	95	3,327	94	94
Oregon	85	88	100	94	2,614	80	83
Pennsylvania	100	100	114	94	3,336	94	94
Rhode Island	100	100	113	94	3,467	94	94
South Carolina	99	99	105	95	2,406	94	94
Tennessee ‡	78	78	91	96	2,930	75	75
Texas	89	89	139	95	3,609	84	84
Utah	100	100	111	94	3,645	94	94
Vermont	90	90	106	95	1,663	85	85
Virginia	100	100	109	95	3,115	95	95
Washington ‡	75	75	85	95	2,423	71	71
West Virginia	99	99	135	96	2,462	95	95
Wisconsin ‡	55	55	63	95	1,427	52	52
Wyoming	100	100	160	95	2,704	95	95
Other Jurisdictions							
District of Columbia	100	100	117	90	2,553	90	90
DDESS [1]	99	99	39	96	1,299	95	95
DoDDS [2]	99	99	91	95	2,850	94	94
Guam	100	100	25	96	1,191	96	96
Virgin Islands	100	100	24	95	707	95	95

‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table A.3 School and student participation rates, grade 8 public schools: By state, 2002

Grade 8	School participation			Student participation		Overall participation rate	
	Student weighted						
	Percentage before substitution	Percentage after substitution	Number of schools participating after substitution	Student weighted percentage	Number of students assessed	Before substitution	After substitution
Nation (Public)	83	84	4,208	91	112,485	76	77
Alabama	80	93	100	93	2,625	75	87
Arizona	93	93	110	88	2,456	82	82
Arkansas	99	99	103	91	2,556	90	90
California ‡	71	71	125	90	3,140	64	64
Connecticut	100	100	104	92	2,707	92	92
Delaware	100	100	35	90	3,903	90	90
Florida	100	100	105	91	2,706	91	91
Georgia	100	100	111	93	3,858	93	93
Hawaii	100	100	55	93	2,745	93	93
Idaho	86	86	80	93	2,455	80	80
Illinois ‡	56	56	106	90	2,416	51	51
Indiana	98	98	101	91	2,586	89	89
Kansas ‡	72	72	84	93	1,898	67	67
Kentucky	96	96	100	94	2,609	90	90
Louisiana	98	98	98	93	2,372	91	91
Maine	94	94	102	92	2,639	86	86
Maryland	93	93	99	90	2,467	84	84
Massachusetts	98	98	104	93	2,679	91	91
Michigan	98	98	104	88	2,450	86	86
Minnesota ‡	66	66	67	91	1,695	60	60
Mississippi	94	94	96	93	2,459	87	87
Missouri	92	96	114	91	2,620	84	88
Montana ‡	76	76	78	94	1,915	71	71
Nebraska	99	99	102	92	2,222	91	91
Nevada	100	100	65	88	2,582	88	88
New Mexico	93	93	91	92	2,389	86	86
New York ‡	71	71	84	88	1,971	63	63
North Carolina	100	100	106	93	2,698	93	93
North Dakota ‡	77	77	112	94	2,051	73	73
Ohio	96	96	94	90	2,337	87	87
Oklahoma	100	100	123	92	2,576	92	92
Oregon ‡	78	78	85	91	1,967	71	71
Pennsylvania	100	100	104	92	2,777	92	92
Rhode Island	100	100	55	89	2,608	89	89
South Carolina	97	97	99	93	2,220	90	90
Tennessee ‡	74	74	82	92	2,074	69	69
Texas	92	92	127	93	3,300	85	85
Utah	100	100	93	92	2,749	92	92
Vermont	91	91	99	92	2,414	84	84
Virginia	100	100	103	92	2,664	92	92
Washington ‡	74	74	80	90	1,879	66	66
West Virginia	92	92	97	92	2,312	85	85
Wisconsin ‡	66	66	75	92	1,814	61	61
Wyoming	100	100	82	92	2,598	92	92
Other Jurisdictions							
American Samoa	100	100	22	96	470	96	96
District of Columbia	100	100	36	85	1,734	85	85
DDESS [1]	99	99	14	96	733	94	94
DoDDS [2]	99	99	55	95	2,166	94	94
Guam	100	100	7	94	1,085	94	94
Virgin Islands	100	100	8	93	579	93	93

‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

District Samples

Results from the 2002 writing assessments will also be reported (on a trial basis) in a forthcoming report on district-level samples of fourth- and eighth-grade students in the large urban school districts that participated in the Trial Urban District Assessment. These large urban school districts are Atlanta, Chicago, Houston, Los Angeles, and New York. The sample of students in the urban school districts represents an augmentation of students who would "normally" be selected as part of state samples. These samples allow reliable subgroup reporting in these districts. Furthermore, all students at "lower" geographic levels are assumed to be part of "higher-level" samples. For example, Houston is one of the urban districts included in the Trial Urban District Assessment. Data from students tested in the Houston sample would be used to report results for Houston, and would also contribute to the Texas estimates and to the national calculations.

Standards for State Sample Participation and Reporting of Results

In carrying out the 2002 state assessment program, the National Center for Education Statistics (NCES) established participation rate standards that jurisdictions were required to meet in order for their results to be reported. NCES also established additional standards that required the annotation of published results for jurisdictions whose sample participation rates were low enough to raise concerns about their representativeness. The NCES guidelines used to report results in the state assessments, and the guidelines for notation when there is some risk of nonresponse bias in the reported results, are presented in this section.

Guideline 1

The publication of NAEP results

The conditions that will result in the publication of a jurisdiction's results are presented below.

Guideline 1 — Publication of Public School Results

A jurisdiction will have its public school results published in the 2002 NAEP writing report card (or in other reports that include all state-level results) if and only if its weighted participation rate for the initial sample of public schools is greater than or equal to 70 percent. Similarly, a jurisdiction will receive a separate NAEP State Report if and only if its weighted participation rate for the initial sample of public schools is greater than or equal to 70 percent.

Discussion: If a jurisdiction's public school participation rate for the initial sample of schools is below 70 percent, there is a substantial possibility that bias will be introduced into the assessment results. This possibility remains even after making statistical adjustments to compensate for school nonparticipation. There remains the likelihood that, in aggregate, the substitute schools are sufficiently dissimilar from the originals they are replacing and represent too great a proportion of the population to discount such a difference. Similarly, the assumptions underlying the use of statistical adjustments to compensate for nonparticipation are likely to be significantly violated if the initial response rate falls below the 70 percent level. Guideline 1 takes this into consideration. This guideline is congruent with current NAGB policy, which requires that data for jurisdictions that do not have a 70 percent before-substitution participation rate be reported "in a different format," and with the Education Information Advisory Committee (EIAC) resolution, which calls for data from such jurisdictions not to be published.

The following guidelines concerning school and student participation rates in the NAEP state assessment program were established to address four significant ways in which nonresponse bias could be introduced into the jurisdiction sample estimates. The four significant ways include overall school nonresponse, strata-specific school nonresponse, overall student nonresponse and strata-specific student nonresponse. Presented on the following pages are the conditions that will result in a jurisdiction's receiving a notation in the 2002 reports. Note that in order for a jurisdiction's results to be published with no notations, that jurisdiction must satisfy all guidelines.

Guideline 2

Reporting school and student participation rates with possible bias due to school nonresponse

Guideline 2—Notation for Overall Public School Participation Rate

A jurisdiction that meets Guideline 1 will receive a notation if its weighted participation rate for the initial sample of public schools was below 85 percent and the weighted public school participation rate after substitution was below 90 percent.

Discussion: For jurisdictions that did not use substitute schools, the participation rates are based on participating schools from the original sample. In these situations, the NCES standards specify weighted school participation rates of at least 85 percent to guard against potential bias due to school nonresponse. Thus the first part of these guidelines, referring to the weighted school participation rate for the initial sample of schools, is in direct accordance with NCES standards.

To help ensure adequate sample representation for each jurisdiction participating in the NAEP 2002 state assessments, NAEP provided substitutes for nonparticipating public schools. For jurisdictions that used substitute schools, the assessment results will be based on the student data from all schools participating from both the original sample and the list of substitutes (unless both an initial school and its substitute eventually participated, in which case only the data from the initial school will be used).

The NCES standards do not explicitly address the use of substitute schools to replace initially selected schools that decide not to participate in the assessment. However, considerable technical consideration was given to this issue. Even though the characteristics of the substitute schools were matched as closely as possible to the characteristics of the initially selected schools, substitution does not entirely eliminate bias due to the nonparticipation of initially selected schools. Thus, for the weighted school participation rates including substitute schools, the guidelines were set at 90 percent.

If a jurisdiction meets either standard (i.e., 85 percent or higher prior to substitution or 90 percent or higher after substitution), there will be no notation for the relevant overall school participation rate.

Guideline 3

Important segments of the jurisdiction's student population that must be adequately represented to avoid possible nonresponse bias

Guideline 3—Notation for Strata-Specific Public School Participation Rates

A jurisdiction that is not already receiving a notation under Guideline 2 will receive a notation if the sample of public schools included a class of schools with similar characteristics that had a weighted participation rate (after substitution) of below 80 percent, and from which the nonparticipating schools together accounted for more than 5 percent of the jurisdiction's total weighted sample of public schools. The classes of schools from each of which a jurisdiction needed minimum school participation levels were determined by degree of urbanization, minority enrollment, and median household income of the area in which the school is located.

Discussion: The NCES standards specify that attention should be given to the representativeness of the sample coverage. Thus, if some important segment of the jurisdiction's population is not adequately represented, it is of concern, regardless of the overall participation rate.

If nonparticipating schools are concentrated within a particular class of schools, the potential for substantial bias remains, even if the overall level of school participation appears to be satisfactory. Nonresponse adjustment cells for public schools have been formed within each jurisdiction, and the schools within each cell are similar with respect to degree of urbanization, minority enrollment, and/or median household income, as appropriate for each jurisdiction.

If the weighted response rate, after substitution, for a single adjustment cell falls below 80 percent, and more than 5 percent (weighted) of the sampled schools are nonparticipants from such a cell, the potential for nonresponse bias is too great. This guideline is based on the NCES standard for stratum-specific school response rates.

Guideline 4

Possible student nonresponse bias

Guideline 4—Notation for Overall Student Participation Rate in Public Schools

A jurisdiction that meets Guideline 1 will receive a notation if the weighted student response rate within participating public schools was below 85 percent.

Discussion: This guideline follows the NCES standard of 85 percent for overall student participation rates. The weighted student participation rate is based on all eligible students from initially selected or substitute schools who participated in the assessment in either an initial session or a make-up session. If the rate falls below 85 percent, the potential for bias due to students' nonresponse is too great.

Guideline 5

Possible nonresponse bias from inadequately represented strata

Guideline 5—Notation for Strata-Specific Student Participation Rates in Public Schools

A jurisdiction that is not already receiving a notation under Guideline 4 will receive a notation if the sampled students within participating public schools included a class of students with similar characteristics that had a weighted student response rate of below 80 percent, and from which the nonresponding students together accounted for more than 5 percent of the jurisdiction's weighted assessable public school student sample. Student groups from which a jurisdiction needed minimum levels of participation were determined by the age of the student, whether or not the student was classified as a student with a disability (SD) or limited English proficient (LEP), and the type of assessment session, as well as school level of urbanization, minority enrollment, and median household income of the area in which the school is located.

Discussion: This guideline addresses the fact that if nonparticipating students are concentrated within a particular class of students, the potential for substantial bias remains, even if the overall student participation level appears to be satisfactory. Student nonresponse adjustment cells have been formed using the school-level nonresponse adjustment cells, together with the student's age and the nature of the assessment session (unmonitored or monitored).

If the weighted response rate for a single adjustment cell falls below 80 percent, and more than 5 percent (weighted) of the invited students who do not participate in the assessment are from such a cell, the potential for nonresponse bias is too great. This guideline is based on the NCES standard for stratum-specific student response rates.

At both the fourth and eighth grades, two states, Illinois and Wisconsin, did not meet the initial public-school participation rate of 70 percent. In addition, one state, Minnesota, did not meet this standard at the eighth grade. Results for these jurisdictions are not included with the findings reported for the state NAEP 2002 writing assessment.

Nine jurisdictions at grade 4 did not meet the second guideline for notation (i.e., the weighted participation rate for the initial sample of schools was below 85 percent and the weighted school participation rate after substitution was below 90 percent): California, Iowa, Kansas, Minnesota, Montana, New York, North Dakota, Tennessee, and Washington. At grade 8, eight jurisdictions did not meet this guideline: California, Kansas, Montana, New York, North Dakota, Oregon, Tennessee, and Washington. Results for each of these jurisdictions at the appropriate grade level are shown with a notation indicating possible bias related to nonresponse.

Students with Disabilities (SD) and/or Limited English Proficient (LEP) Students

It is NAEP's intent to assess all selected students from the target population. Therefore, every effort is made to ensure that all selected students who are capable of participating in the assessment are assessed. Some students sampled for participation in NAEP can be excluded from the sample according to carefully defined criteria. These criteria communicate a presumption of inclusion except under special circumstances. According to these criteria, students who had an Individualized Education Program (IEP) or were protected under Section 504 of the Rehabilitation Act of 1973[4] were to be included in the NAEP assessment except in the following cases:

- the school's IEP team determined that the student could not participate,
- the student's cognitive functioning was so severely impaired that she or he could not participate, or
- the student's IEP required that the student be tested with an accommodation or adaptation that NAEP does not allow and that the student could not demonstrate his or her knowledge without that accommodation.

All LEP students who received academic instruction in English for three years or more were to be included in the assessment. Those LEP students who received instruction in English for fewer than three years were to be included unless school staff judged them to be incapable of participating in the assessment in English.

Participation of SD and/or LEP Students in the NAEP Samples

Testing all sampled students is the best way for NAEP to ensure that the statistics generated by the assessment are as representative as possible of the performance of the entire national population and the populations of participating jurisdictions. However, all groups of students include certain proportions that cannot be tested in large-scale assessments (such as students who have profound mental disabilities) or who can only be tested through the use of testing accommodations such as extra time, one-on-one administration, or use of

[4] Section 504 of the Rehabilitation Act of 1973 is a civil rights law designed to prohibit discrimination on the basis of disability in programs and activities, including education, that receive federal financial assistance.

magnifying equipment. Some students with disabilities and some LEP students cannot show on a test what they know and can do unless they are provided with accommodations. When such accommodations are not allowed, students requiring such adjustments are often excluded from large-scale assessments such as NAEP. This phenomenon has become more common in the last decade and gained momentum with the passage of the 1997 Individuals with Disabilities Education Act (IDEA), which led schools and states to identify increasing proportions of students as needing accommodations on assessments in order to best show what they know and can do.[5] Furthermore, section 504 of the Rehabilitation Act of 1973 requires that, when students with disabilities are tested, schools must provide them with appropriate accommodations so that the test results accurately reflect students' achievement. In addition, as the proportion of limited English proficient students in the population has increased, some states have started offering accommodations, such as translations of assessments or the use of bilingual dictionaries as part of assessments.

Before 1996, NAEP did not allow any testing under nonstandard conditions (i.e., accommodations were not permitted). At that time, NAEP samples were able to include almost all sampled students in standard assessment sessions. However, as the influence of IDEA grew more widespread, the failure to provide accommodations led to increasing levels of exclusion in the assessment. Such increases posed two threats to the program: 1) they threatened the stability of trend lines (because excluding more students in one year than the next might lead to apparent rather than real gains), and 2) they made NAEP samples less than optimally representative of target populations.

The reporting samples in the 1998 and 2002 writing assessments used these criteria with provisions made for accommodations. Students with disabilities or limited English proficient students were given accommodations that matched as closely as possible those provided to them in other testing situations by their schools or instructors (most frequently, extended time for responding). All the scale score and achievement level information in this report, then, is based on a student sample that includes students who were provided with accommodations. The responses of students assessed with accommodations were evaluated according to the same criteria as those of students assessed without accommodations.

In order to make it possible to evaluate both the impact of increasing exclusion rates in some jurisdictions and differences between jurisdictions, complete data on exclusion in all years are included in this appendix. Since the exclusion rates may affect trend measurement within a jurisdiction, readers should consider the magnitude of exclusion rate changes when interpreting score changes in jurisdictions. In addition, different rates of exclusion may influence the meaning of state comparisons. Thus, exclusion data should be reviewed in this context as well.

[5] Office of Special Education Programs. (1997). *Nineteenth Annual Report to Congress on the Implementation of the Individuals With Disabilities Education Act*. Washington, DC: U. S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Percentages of SD and/or LEP students for the 1998 and 2002 national sample are presented in table A.4. The data in this table include the percentage of students identified as SD and/or LEP, the percentage of students excluded, the percentage of assessed SD and/or LEP students, the percentage assessed without accommodations, and the percentage assessed with accommodations. Table A.4 also includes similar data for SD students only and LEP students only. Tables A.5 and A.6 show similar information by jurisdiction for grade 4 and grade 8. Table A.5 shows 2002 data only since the 1998 state assessments were administered only at grade 8.

In the 2002 national sample, 5 percent of students at grade 4, 4 percent of students at grade 8, and 3 percent of students at grade 12 were excluded from the assessment (See table A.4). Across the various jurisdictions that participated in the 2002 state assessment, the percentage of students excluded ranged from 2 to 10 percent at grade 4 (see table A.5) and from 1 to 8 percent at grade 8 (see table A.6).

Table A.4 Students with disabilities and/or limited English proficient students identified, excluded, and assessed, grades 4, 8, and 12: 1998 and 2002

	1998		2002	
	Number of students	Weighted percentage of all students sampled	Number of students	Weighted percentage of all students sampled
Grade 4				
SD[1] and/or LEP[2] students				
Identified	3,621	15	26,998	19
Excluded	1,450	5	7,608	5
Assessed	2,171	10	19,390	14
Without accommodations	1,425	6	11,281	9
With accommodations	746	4	8,109	5
SD[1] students				
Identified	2,192	11	19,052	12
Excluded	806	4	5,603	4
Assessed	1,386	7	13,449	8
Without accommodations	744	4	6,153	4
With accommodations	642	3	7,296	4
LEP[2] students				
Identified	1,582	4	9,923	8
Excluded	753	2	2,878	2
Assessed	829	2	7,045	7
Without accommodations	709	2	5,777	6
With accommodations	120	#	1,268	1
Grade 8				
SD[1] and/or LEP[2] students				
Identified	2,935	13	20,516	17
Excluded	877	4	5,012	4
Assessed	2,058	9	15,504	13
Without accommodations	1,380	6	8,877	8
With accommodations	678	3	6,627	5
SD[1] students				
Identified	2,139	10	16,420	12
Excluded	672	3	3,958	3
Assessed	1,467	7	12,462	9
Without accommodations	863	5	6,250	5
With accommodations	604	3	6,212	5
LEP[2] students				
Identified	924	3	5,526	6
Excluded	273	1	1,554	1
Assessed	651	2	3,972	4
Without accommodations	561	2	3,211	4
With accommodations	90	#	761	1

See footnotes at end of table. ▶

Table A.4 Students with disabilities and/or limited English proficient students identified, excluded, and assessed, grades 4, 8, and 12: 1998 and 2002—Continued

Grade 12	1998 Number of students	1998 Weighted percentage of all students sampled	2002 Number of students	2002 Weighted percentage of all students sampled
SD[1] and/or LEP[2] students				
Identified	1,975	8	2,120	11
Excluded	658	2	754	3
Assessed	1,317	6	1,366	8
Without accommodations	991	5	919	6
With accommodations	326	1	447	3
SD[1] students				
Identified	1,375	6	1,654	9
Excluded	566	2	674	3
Assessed	809	4	980	6
Without accommodations	536	3	574	4
With accommodations	273	1	406	3
LEP[2] students				
Identified	654	2	591	3
Excluded	122	#	146	1
Assessed	532	2	445	2
Without accommodations	474	2	389	2
With accommodations	58	#	56	#

Percentage rounds to zero.
[1] Students with disabilities.
[2] Limited English proficient students.
NOTE: Within each grade level, the combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions, because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions, but counted only once in the top portion.
Within each portion of the table, percentages may not add to totals due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table A.5 **Percentage of students with disabilities and/or limited English proficient students identified, excluded, and assessed, grade 4 public schools: By state, 2002**

Grade 4	2002					
	SD[1] and/or LEP[2] students					
	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations	All students assessed without accommodations
Nation (Public)	20	5	15	10	5	89
Alabama	14	2	12	9	3	95
Arizona	28	6	22	19	3	90
Arkansas	15	3	12	7	5	92
California	35	4	30	29	2	94
Connecticut	16	5	11	5	6	89
Delaware	17	6	11	4	8	86
Florida	24	6	19	9	10	84
Georgia	14	3	11	5	5	91
Hawaii	18	4	15	8	7	90
Idaho	18	2	16	10	6	92
Illinois	20	6	14	7	7	87
Indiana	13	4	10	6	3	93
Iowa	16	5	11	3	8	87
Kansas	20	3	17	6	10	86
Kentucky	11	6	5	3	2	92
Louisiana	19	4	15	4	12	84
Maine	18	5	13	6	7	88
Maryland	15	7	7	6	2	91
Massachusetts	19	5	14	3	11	84
Michigan	13	5	8	5	3	91
Minnesota	19	4	14	9	5	91
Mississippi	7	4	3	2	1	95
Missouri	16	5	11	4	7	88
Montana	14	4	10	4	5	91
Nebraska	19	3	16	9	7	90
Nevada	26	8	19	13	6	87
New Mexico	37	7	30	21	8	84
New York	19	7	12	4	8	85
North Carolina	19	7	12	3	9	84
North Dakota	17	3	13	8	6	91
Ohio	12	7	5	3	2	91
Oklahoma	19	3	15	9	6	91
Oregon	24	6	17	12	6	88
Pennsylvania	14	4	10	4	6	91
Rhode Island	23	4	19	8	11	85
South Carolina	17	5	12	9	4	92
Tennessee	15	3	12	9	3	94
Texas	26	10	16	13	2	87
Utah	20	4	17	11	6	90
Vermont	15	5	11	3	8	88
Virginia	19	6	13	5	8	86
Washington	14	3	11	6	5	92
West Virginia	15	5	10	4	6	89
Wisconsin	19	7	11	5	6	86
Wyoming	16	2	14	6	8	90
Other Jurisdictions						
District of Columbia	19	6	12	6	6	87
DDESS [3]	17	3	14	8	6	91
DoDDS [4]	16	3	13	9	4	93
Guam	38	4	34	27	7	90
Virgin Islands	8	4	5	3	1	95

[1] Students with disabilities
[2] Limited English proficient students
[3] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[4] Department of Defense Dependents Schools (Overseas).
NOTE: Percentages may not add to totals due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table A.6 Percentage of students with disabilities and/or limited English proficient students identified, excluded, and assessed, grade 8 public schools: By state, 1998 and 2002

Grade 8	1998						2002					
	SD[1] and/or LEP[2] students					All students assessed without accommodations	SD[1] and/or LEP[2] students					All students assessed without accommodations
	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations		Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations	
Nation (Public)	14	4	10	7	3	93	18	4	14	8	5	90
Alabama	12	6	6	5	1	93	15	3	12	11	1	96
Arizona	17	5	12	10	2	92	22	5	17	14	3	92
Arkansas	13	6	7	5	1	93	17	3	14	9	5	92
California	23	6	17	15	2	92	27	3	24	20	3	93
Colorado	13	4	9	6	3	93	—	—	—	—	—	—
Connecticut	15	7	8	5	3	90	17	4	13	7	6	90
Delaware	14	3	11	8	3	94	15	5	11	2	8	87
Florida	16	5	11	9	2	93	20	4	16	7	10	87
Georgia	11	5	7	4	2	93	13	3	10	5	5	92
Hawaii	15	4	11	8	3	93	21	3	18	11	7	90
Idaho	—	—	—	—	—	—	14	2	13	8	4	94
Illinois	12	4	8	6	2	94	18	3	14	8	7	90
Indiana	—	—	—	—	—	—	13	3	10	7	3	94
Kansas	—	—	—	—	—	—	16	3	13	6	7	90
Kentucky	10	2	7	3	4	93	11	4	8	4	3	93
Louisiana	13	5	8	3	5	90	16	4	12	4	8	88
Maine	14	5	8	5	3	92	18	2	16	8	8	90
Maryland	13	2	11	4	7	91	16	4	12	9	3	93
Massachusetts	17	5	12	7	5	90	20	3	16	7	10	87
Michigan	—	—	—	—	—	—	14	5	9	4	4	90
Minnesota	14	3	11	8	3	94	17	3	14	9	5	92
Mississippi	9	5	5	4	1	94	10	5	5	3	2	93
Missouri	13	3	10	6	4	93	16	3	13	4	9	88
Montana	11	2	9	6	2	95	13	2	12	7	4	94
Nebraska	—	—	—	—	—	—	17	4	12	7	5	91
Nevada	16	6	10	8	3	91	21	4	16	12	5	91
New Mexico	23	6	17	14	3	90	32	5	27	17	10	85
New York	15	5	9	3	6	89	20	6	14	5	9	85
North Carolina	14	4	10	4	6	89	17	5	12	4	9	87
North Dakota	—	—	—	—	—	—	15	1	14	8	6	93
Ohio	—	—	—	—	—	—	12	6	7	4	2	92
Oklahoma	13	9	5	4	1	90	16	2	14	9	4	93
Oregon	15	3	12	9	3	94	18	4	14	11	3	93
Pennsylvania	—	—	—	—	—	—	14	2	12	4	8	90
Rhode Island	17	4	13	10	3	93	22	3	18	9	10	87
South Carolina	12	5	7	5	2	93	15	5	10	6	4	91
Tennessee	13	4	9	8	1	95	14	3	12	10	2	95
Texas	19	6	13	10	2	92	19	7	13	11	2	92
Utah	10	4	6	5	1	95	17	3	14	9	4	93
Vermont	—	—	—	—	—	—	17	4	14	6	7	89
Virginia	14	4	9	6	3	93	18	6	12	5	7	87
Washington	13	4	9	7	3	94	15	3	11	6	5	91
West Virginia	14	5	9	5	3	92	18	4	14	5	9	86
Wisconsin	11	4	7	4	3	93	17	4	13	4	9	87
Wyoming	9	2	7	5	2	96	15	2	13	6	7	91
Other Jurisdictions												
American Samoa	—	—	—	—	—	—	22	7	15	9	6	87
District of Columbia	13	6	7	6	1	92	21	6	15	5	10	84
DDESS [3]	10	3	7	4	3	94	15	3	12	5	7	90
DoDDS [4]	7	1	6	4	2	97	10	1	8	6	3	96
Guam	—	—	—	—	—	—	31	1	30	27	3	95
Virgin Islands	8	8	#	#	0	92	10	8	2	2	#	92

— Indicates that the jurisdiction did not participate.
Percentage rounds to zero.
[1] Students with disabilities [2] Limited English proficient students
[3] Department of Defense Domestic Dependent Elementary and Secondary Schools. [4] Department of Defense Dependents Schools (Overseas).
NOTE: Percentages may not add to totals due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Investigating the Potential Effects of Exclusion Rates on Assessment Results

Variation in the rates of exclusion of students with disabilities and limited English proficient students introduces validity concerns for comparisons over time or between jurisdictions. The essential problem is the differential representativeness of samples, which could impact the comparability of cross-state comparisons within a given year and state trends across years. Since students with disabilities or limited English proficient students tend to score below average on assessments, excluding students with special needs may increase a jurisdiction's scores. Conversely, including more of these students might depress score gains. In 2002, exclusion rates varied among jurisdictions. In addition, cases of both increases and decreases in exclusion rates occurred between 1998 and 2002, making comparisons over time within jurisdictions complex to interpret. Tables A.5 and A.6 on the preceding pages display the rates of exclusion in each jurisdiction for grade 4 in 2002 and for grade 8 in 2002 and 1998, respectively.

As shown in table A.5, of the 48 jurisdictions that assessed writing at grade 4 in 2002, all jurisdictions except Texas had exclusion rates of less than 10 percent, and more than two-thirds had exclusion rates of less than six percent. Table A.6 displays the comparable data for grade 8. In 2002, all jurisdictions had exclusion rates less than 9 percent and about three-quarters had exclusion rates of less than five percent. Exclusion rates in grade 8 increased from 1998 to 2002 in eight jurisdictions, with an average increase of 1.5 percentage points.

At least two factors contributed to the variability in exclusion rates across states. One factor is that the percentage of students who are *identified* as having disabilities or limited proficiency in English varies across jurisdictions and over time. Reasons for this variation include: 1) lack of standardized criteria for defining students as having specific disabilities or as being limited in their English proficiency; 2) changes or differences in policy and practices regarding implementation of the Individuals with Disabilities Education Act (IDEA); and 3) population shifts in the percentage of students classified as limited English proficient and, to a lesser extent, as students with disabilities.

Another factor is that some SD and/or LEP students are excluded because they are so severely disabled or lacking in English language skills that no accommodation would be sufficient to enable them to participate meaningfully.

With regard to cross-state comparisons, the correlations between rates of exclusion and average writing scores were not found to be significant at grade 4 (.18). Because exclusion is not significantly related to scores, states that exclude more SD and/or LEP students would not have an advantage over other states. At eighth grade, the correlation between rates of exclusion and average writing scores was not significant (−.27) in 1998; but was significant (−.33) in 2002. Since the direction of the correlation is negative, states that excluded more eighth grade students in 2002 would be disadvantaged in comparison with other states that excluded fewer eighth grade students.

With regard to state trends, the correlations between changes in the rate of exclusion of students with special needs and average writing score gains from 1998 to 2002 were found to be moderate (.51 at grade 8). While there was a moderate tendency for an increase in exclusion rates to be associated with an increase in average scale scores, exclusion increases do not explain the entirety of score gains.

Because the representativeness of samples is ultimately a validity issue, NCES has commissioned studies of the impact of assessment accommodations on overall scores. NCES has also investigated scenarios for estimating what the average scores might have been had the excluded students been assessed. Two alternative statistical scenarios have been proposed, based on different hypotheses about how excluded students might have performed. Combined with the actual performance of students who were assessed, these scenarios produce results for the full population (that is, including estimates for excluded students) in each jurisdiction and each assessment year. These techniques provide some indication as to which statements about trend gains or losses *might* be changed if exclusion rates were zero in both assessment years and if the hypotheses about the performance of missing students are correct.

One scenario was developed by Donald McLaughlin of American Institutes for Research, and predicts what the performance of excluded SD and/or LEP students might have been had these students been tested. The basic assumption underlying this approach is that these students would have performed as well as included SD and/or LEP students with similar disabilities, level of English proficiency, and background characteristics.[6]

The other scenario was developed by Al Beaton of Boston College and similarly makes an assumption about what the performance of excluded SD/LEP students might have been had they been tested. The idea of Beaton's scenario is to calculate median, rather than average scores. A 'median' is the score reached or exceeded by fifty percent of the student population. This statistic is not influenced by extreme values. Beaton's assumption is that all SD/LEP students would score below *Basic* or below the median of the group being analyzed. This assumption lowers the median score for every group.

The methods used to construct the scenarios are still under development. NCES is continuing research into different procedures for reducing the percentages of students excluded from NAEP. In addition, NCES will continue to evaluate the potential impact of changes in exclusion rates on score gains. More detailed information on the scenarios will be available at the NAEP web site at *http://www.nces.ed.gov/nationsreportcard*. The scenarios illustrate the potential impact of reasonable hypotheses about the performance of excluded students on score gains in the jurisdictions that participated in both 1998 and 2002 and should not be interpreted as official results.

[6] Because students with very severe levels of disability and students with little or no proficiency in English are not assessed in NAEP, ability estimates for students with those characteristics may be overestimated.

Types of Accommodations Permitted

Table A.7 displays the percentages of SD and/or LEP students assessed with the variety of available accommodations. It should be noted that students assessed with accommodations typically received some combination of accommodations. The percentages presented in the table reflect only the primary accommodation provided. For example, students assessed in small groups (as compared with standard NAEP sessions of about 30 students) usually received extended time. In one-on-one administrations, students often received assistance in recording answers and were afforded extra time. Extended time was considered the primary accommodation only when it was the sole accommodation provided.

Table A.7 Students with disabilities and/or limited English proficient students assessed with accommodations, by type of primary accommodation, grades 4, 8, and 12 public and nonpublic schools: 1998 and 2002

	Weighted percentage of students sampled					
	Grade 4		Grade 8		Grade 12	
	1998	2002	1998	2002	1998	2002
SD[1] and/or LEP[2] students						
Bilingual dictionary	0.01	0.02	0.04	0.11	0.02	0.09
Large-print book	0.01	0.03	0.02	0.03	0.01	0.01
Extended time	0.76	1.52	0.82	1.84	0.45	1.35
Read aloud	0.28	0.31	0.08	0.27	0.04	0.16
Small group	2.31	3.08	1.61	2.62	0.67	1.07
One-on-one	0.23	0.13	0.12	0.11	0.07	0.06
Scribe/computer	0.17	0.02	0.05	0.02	0.04	0.02
Other	0.02	0.02	0.02	0.05	0.05	0.02
SD[1] students only						
Bilingual dictionary	#	#	#	0.01	#	#
Large-print book	0.01	0.03	0.02	0.03	0.01	0.01
Extended time	0.65	1.21	0.71	1.65	0.35	1.26
Read aloud	0.25	0.29	0.06	0.24	0.03	0.15
Small group	2.17	2.77	1.58	2.52	0.65	1.05
One-on-one	0.22	0.13	0.11	0.11	0.07	0.06
Scribe/computer	0.17	0.02	0.05	0.02	0.04	0.02
Other	0.02	0.02	0.02	0.05	0.05	0.02
LEP[2] students only						
Bilingual dictionary	0.01	0.02	0.04	0.11	0.02	0.09
Large-print book	#	#	#	#	#	#
Extended time	0.13	0.43	0.11	0.34	0.10	0.13
Read aloud	0.05	0.03	0.03	0.04	#	#
Small group	0.17	0.46	0.06	0.24	0.04	0.05
One-on-one	0.01	0.01	0.01	0.01	#	#
Scribe/computer	#	#	#	#	#	#
Other	#	0.01	#	0.01	#	#

Percentage rounds to less than 0.01.
[1] Students with disabilities.
[2] Limited English proficient students.
NOTE: The combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions but counted only once in the top portion.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Data Collection and Scoring

The writing assessment was conducted from January to March 2002. Data collection for the 2002 assessment was conducted by trained field staff from Westat.

Materials from the NAEP 2002 writing assessment were shipped to Pearson, where trained staff evaluated the responses to the writing tasks using scoring rubrics or guides prepared by ETS. All the writing tasks were evaluated according to six-level scoring guides. At each grade, scoring guides were developed for each of the three types of tasks: narrative, informative, and persuasive.

Specialists in writing who are highly experienced in teaching and/or assessing writing trained the professional raters who evaluated the student responses. The trainers received intensive training together that included reading a manual that explained how to use the scoring guides and the processes for training and checking raters. For each task, the trainer, in consultation with other trainers or assessment specialists, chose numerous sample responses to present to raters and prepared notes on how the scoring guide applied to the particular task. The sample responses helped raters become accustomed to the variety of responses the task elicited before they began rating the student responses. Raters had to pass a qualifying test before they could evaluate student responses: they had to agree with at least 70 percent of the ratings (to a set of ten student responses) that were given beforehand by their trainer.

In order to determine interrater reliability of scoring, a specified percentage of responses was read twice: two raters read 6 percent of the responses at grades 4 and 8 (grades at which the assessment data was collected from the combined sample), and 25 percent of responses at grade 12.

For the national and state writing assessments, 608,269 responses to writing tasks were scored. This number includes rescoring to monitor interrater reliability. The within-year average percentage of exact agreement of ratings on the six-level scoring guides for the 2002 reliability samples was 83 percent at fourth grade, 82 percent at eighth grade, and 78 percent at twelfth grade.

Data Analysis and IRT Scaling

Subsequent to the professional scoring, all information was transcribed into the NAEP database at ETS. Each processing activity was conducted with rigorous quality control. After the assessment information was compiled in the database, the data were weighted according to the population structure. The weighting for the combined sample reflected the probability of selection for each student as a result of the sampling design, adjusted for nonresponse.[7]

[7] Weighting procedures are described more fully in the "Weighting and Variance Estimation" section later in this document. Additional information about the use of weighting procedures will be included in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

Analyses were then conducted to determine the percentages of students who wrote responses to each writing task at each level on the scoring guide and who provided various responses to each background question. In calculating response percentages for each task, only students classified as having been presented the question were included in the denominator of the statistic. Students whose papers were blank or whose responses were judged to be off topic were similarly excluded from the calculation of the scale.

Item Response Theory (IRT) was used to estimate average writing scale scores for the nation, for various subgroups of interest within the nation, and for the states and other jurisdictions. IRT models the probability of answering a question in a certain way as a mathematical function of proficiency or skill. The main purpose of IRT analysis is to provide a common scale on which performance can be compared among groups such as those defined by characteristics, including gender and race/ethnicity.

The results for the 2002 writing assessments are presented on the NAEP writing scales. In 2002, a scale ranging from 0 to 300 was computed to report performance at each grade level. The scale summarizes student performance across all three purposes for writing (narrative, informative, and persuasive) in the assessment.

In producing the writing scale, an IRT model was used. The writing tasks (all rated according to six-level scoring guides) were scaled by use of a generalized partial-credit (GPC) model.[8] 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 that are used for more complex constructed-response questions such as writing tasks.[9]

Because of the PBIB spiraling design used by NAEP, students do not receive enough writing tasks to provide reliable information about individual performance. Traditional test scores for individual students, even those based on IRT, would result in misleading estimates of population characteristics, such as subgroup means and percentages of students at or above a certain scale score level. However, it is NAEP's goal to estimate these population characteristics. NAEP's objectives can be achieved with methodologies that produce estimates of the population-level parameters directly, without the intermediary computation of estimates of individuals.[10] This is accomplished using marginal estimation scaling model techniques for latent variables. Under the assumptions of the scaling models, these population estimates will be consistent in the sense that the estimates approach the model-based population

[8] Muraki, E. (1992). A Generalized Partial Credit Model: Application of an EM Algorithm. *Applied Psychological Measurement*, (*16)*2, 159–176.

[9] More detailed information regarding the IRT analyses used in NAEP will be included in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

[10] Mislevy, R. J., and Sheehan, K. M. (1987). Marginal Estimation Procedures. In A. E. Beaton (Ed.) Implementing the New Design: The NAEP 1983–1984 Technical Report (Report No: 15-TR-20). Princeton, NJ: Educational Testing Service.

values as the sample size increases. This would not be the case for population estimates obtained by aggregating optimal estimates of individual performance.[11]

Item Mapping Procedures

Item mapping is a procedure by which a rating on a writing task (such as "Sufficient" or better) is associated with a certain point on the 0–300 writing scale. The item maps for writing are presented at the end of chapter 4. For example, the "Sufficient" rating for a given writing task will map onto the scale at 150 if students with an average scale score of at least 150 have a good chance of earning a rating of "Sufficient" or better. It is not clear how to define "a good chance" in terms of the probability, expressed as a percentage, that a given student will respond to an item at the score level designated. A response-probability convention has to be adopted that will divide those students who have a higher probability of success from those who have a lower probability. Which response-probability convention is adopted largely determines where ratings on writing tasks will map onto the writing scale. A lower-boundary convention maps the ratings on writing tasks to lower points on the scale, and a higher-boundary convention maps the same ratings on tasks to higher points on the scale. The underlying distribution of writing 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 tasks on the writing scale.

There is no obvious choice of a point along the probability scale that is clearly superior to any other point. On one hand, if the convention were set with a boundary at 50 percent, those above the boundary would be more likely to score at a particular rating (or higher) on the task than not, while those below the boundary would be more likely to receive a lower rating. Although this convention has some intuitive appeal, it was rejected on the grounds that having a 50:50 chance of getting a particular rating shows an insufficient degree of mastery. On the other hand, if the convention were set with a boundary at 80 percent, students above the criterion would have a high probability of receiving a given rating or higher. However, many students below this criterion may possess substantial writing ability that would be ignored by such a stringent criterion. In particular, those with a 50–80 percent probability of receiving a particular rating (or higher) would be more likely to receive that rating than not, yet would not be in the group described as "able to achieve" that level of performance on the task.

[11] For theoretical and empirical justification of the procedures employed, see Mislevy, R. J. (1988). Randomization-Based Inferences About Latent Variables From Complex Samples. *Psychometrika, (56)*2, 177–196.

In a compromise between the 50 percent and the 80 percent conventions, NAEP has adopted a response probability convention of 65 percent for constructed-response questions such as writing tasks. This probability convention was established, in part, based on an intuitive judgment that it would provide the best picture of students' writing ability.

Some additional support for this convention was provided by Huynh.[12] He examined the IRT information provided by items, according to the IRT model used in scaling NAEP questions. Following Bock, Huynh decomposed the item information into that provided by a correct response [P(q) I(q)] and that provided by an incorrect response[(1– P(q)) I(q)].[13] Huynh showed that the item information provided by a correct response to a constructed-response item is maximized at the point along the writing scale at which the probability of a correct response is two-thirds. It should be noted, however, that maximizing the item information I(q), rather than the information provided by a correct response [P(q) I(q)], would imply an item mapping criterion closer to 50 percent.

Weighting and Variance Estimation

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

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

Not only must appropriate estimates of population characteristics be derived, but appropriate measures of the degree of uncertainty must be obtained for those statistics. Two components of uncertainty are accounted for in the variability of statistics based on student ability: (a) the uncertainty due to sampling only a relatively small number of students, and (b) the uncertainty due to sampling only a relatively small number of cognitive questions (in this case, writing tasks). The first component accounts for the variability associated with the estimated percentages of students who had certain background characteristics or who had a certain rating for their responses to a task.

Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate.

[12] Huynh, H. (1998). On Score Locations of Binary and Partial Credit Items and Their Application to Item Mapping and Criterion-Referenced Interpretation. *Journal of Educational and Behavioral Statistics, 23*(1), 35–56.

[13] Bock, R. D. (1972). Estimating Item Parameters and Latent Ability When Responses are Scored in Two or More Latent Categories. *Psychometrika, 37,* 29–51.

NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any student information that can be observed without error. However, because each student typically responds to only two writing tasks, the scale score for any single student would be imprecise. In this case, NAEP's marginal estimation methodology can be used to describe the performance of groups and subgroups of students. The estimate of the variance of the students' posterior scale score distributions (which reflect the imprecision due to lack of measurement accuracy) is computed. This component of variability is then included in the standard errors of NAEP scale scores.[14]

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 estimation of standard errors may be quite large. In such cases, the standard errors—and any confidence intervals or significance tests involving these standard errors—should be interpreted cautiously. Additional details concerning procedures for identifying such standard errors will be discussed in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

The reader is reminded that, as with findings from all surveys, NAEP results are subject to other kinds of error, including the effects of imperfect adjustment for student and school nonresponse and 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 background information; mistakes in recording, coding, or scoring data; and other errors in collecting, processing, sampling, and estimating missing data. The extent of nonsampling errors is difficult to estimate and, because of 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

Because the percentages of students in these subpopulations and their average scale scores are based on samples rather than on the entire population of fourth-, eighth-, or twelfth-graders in the nation or a jurisdiction, the numbers reported are estimates. As such, they are subject to a measure of uncertainty, reflected in the standard error of the estimate. When the estimated percentages or average scale scores of certain groups are compared, the standard error should be taken into account, and observed similarities or differences should not be relied on solely. Therefore, the comparisons discussed in this report are based on statistical tests that consider the estimated standard errors of those statistics and the magnitude of the difference among the averages or percentages.

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

For the data presented in this report, all the estimates have corresponding estimated standard errors. For example, table A.8 shows the average scale score for the NAEP 1998 and 2002 national writing assessments, and table A.9 shows the percentage of students within each achievement level range and at or above achievement levels. In both tables, estimated standard errors appear in parentheses next to each estimated scale score or percentage. Additional examples of estimated standard errors corresponding with results included in this report are presented in tables A.10, A.11, and A.12. For the estimated standard errors corresponding to other data in this report, the reader can go to the data tool on the NCES web site at *http://nces.ed.gov/nationsreportcard/naepdata*.

Using confidence intervals based on the standard errors provides a way to take into account the uncertainty associated with sample estimates and to make inferences about the population averages and percentages in a manner that reflects that uncertainty. An estimated sample average scale score plus or minus 1.96 standard errors approximates a 95 percent confidence interval for the corresponding population quantity. This statement means that one can conclude with an approximately 95 percent level of confidence that the average performance of the entire population of interest (e.g., all fourth-grade students in public and nonpublic schools) is within plus or minus 1.96 standard errors of the sample average.

For example, suppose that the average writing scale score of the students in a particular group was 162 with an estimated standard error of 1.2. An approximately 95 percent confidence interval for the population quantity would be as follows:

Average $\pm$ 1.96 standard errors

$162 \pm 1.96 \quad 1.2$

$162 \quad 2.4$

(159.6, 164.4)

Thus, one can conclude with a 95 percent level of confidence that the average scale score for the entire population of students in that group is between 159.6 and 164.4. It should be noted that this example and the examples in the following sections are illustrative. More precise estimates carried out to one or more decimal places are used in the actual analyses.

Similar confidence intervals can be constructed for percentages, if the percentages are not extremely large or extremely small. Extreme percentages should be interpreted with caution. Adding or subtracting the standard errors associated with extreme percentages could cause the confidence interval to exceed 100 percent or fall below 0 percent, resulting in numbers that are not meaningful. A more complete discussion of extreme percentages will appear in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

Table A.8 Average writing scale scores and standard errors, grades 4, 8, and 12: 1998 and 2002

	1998	2002
Grade 4	150 (0.7) *	154 (0.4)
Grade 8	150 (0.6) *	153 (0.5)
Grade 12	150 (0.7)	148 (0.8)

* Significantly different from 2002.
NOTE: Standard errors of the estimated scale scores appear in parentheses.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table A.9 Percentage of students and standard errors by writing achievement level, grades 4, 8, and 12: 1998 and 2002

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4							
	1998	16 (0.4) *	61 (0.6) *	22 (0.7) *	1 (0.2) *	84 (0.4) *	23 (0.8) *
	2002	14 (0.4)	58 (0.4)	26 (0.4)	2 (0.1)	86 (0.4)	28 (0.4)
Grade 8							
	1998	16 (0.5)	58 (0.5) *	25 (0.7) *	1 (0.1) *	84 (0.5)	27 (0.7) *
	2002	15 (0.4)	54 (0.5)	29 (0.5)	2 (0.1)	85 (0.4)	31 (0.6)
Grade 12							
	1998	22 (0.7) *	57 (0.7) *	21 (0.7)	1 (0.1) *	78 (0.7) *	22 (0.7)
	2002	26 (0.7)	51 (0.7)	22 (0.7)	2 (0.2)	74 (0.7)	24 (0.8)

* Significantly different from 2002.
NOTE: Standard errors of the estimated percentages appear in parentheses.
Percentages within each writing achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table A.10 Average writing scale scores and standard errors, by student eligibility for free/reduced-price school lunch and race/ethnicity , grades 4, 8, and 12: 2002

	Eligible	Not eligible	Information not available
Grade 4			
Total	141 (0.8)	163 (0.5)	161 (1.5)
White	147 (0.5)	165 (0.5)	166 (1.2)
Black	136 (0.8)	150 (1.2)	145 (2.0)
Hispanic	137 (2.2)	155 (1.4)	147 (3.4)
Asian/Pacific Islander	155 (2.7)	173 (1.9)	172 (3.7)
American Indian/Alaska Native	132 (2.2)	151 (3.0)	143 (4.7)
Grade 8			
Total	136 (0.5)	162 (0.7)	161 (1.5)
White	144 (0.7)	164 (0.7)	168 (1.6)
Black	129 (0.7)	145 (1.1)	142 (2.1)
Hispanic	131 (1.1)	149 (1.5)	143 (2.0)
Asian/Pacific Islander	144 (2.6)	170 (2.9)	166 (5.5)
American Indian/Alaska Native	127 (3.8)	151 (3.5)	135 (5.0) !
Grade 12			
Total	132 (1.4)	152 (1.0)	156 (1.5)
White	139 (1.9)	154 (1.0)	159 (1.5)
Black	123 (1.5)	134 (2.0)	137 (3.1)
Hispanic	130 (1.6)	139 (2.2)	144 (4.1)
Asian/Pacific Islander	134 (3.1)	155 (3.3)	161 (5.6) !
American Indian/Alaska Native	*** (***)	*** (***)	*** (***)

! The nature of the sample does not allow accurate determination of the variability of the statistic.
***(***) Quality control activities and special analysis raised concerns about the accuracy and precision of grade 12 American Indian data. As a result, they are omitted from this report.
NOTE: Standard errors of the estimated scale scores appear in parentheses.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table A.11 Average writing scale scores and standard errors, grade 8 public schools: By state, 1998 and 2002

Grade 8	1998	2002
Nation (Public) [1]	148 (0.6) *	152 (0.6)
Alabama	144 (1.4)	142 (1.5)
Arizona	143 (1.5)	141 (1.6)
Arkansas	137 (1.2) *,**	142 (1.3)
California ‡	141 (1.8)	144 (1.8)
Colorado	151 (1.3)	—
Connecticut	165 (1.4)	164 (1.5)
Delaware	144 (1.4) *,**	159 (0.6)
Florida	142 (1.2) *,**	154 (1.6)
Georgia	146 (1.3)	147 (1.4)
Hawaii	135 (1.0)	138 (0.8)
Idaho	—	151 (1.3)
Indiana	—	150 (1.5)
Kansas ‡	—	155 (1.5)
Kentucky	146 (1.5)	149 (1.1)
Louisiana	136 (1.4) *,**	142 (1.6)
Maine	155 (1.5)	157 (1.2)
Maryland	147 (1.5) *,**	157 (1.5)
Massachusetts	155 (1.7) *,**	163 (1.5)
Michigan	—	147 (1.6)
Minnesota ‡	148 (1.9)	—
Mississippi	134 (1.3) *,**	141 (1.1)
Missouri	142 (1.4) *,**	151 (1.2)
Montana ‡	150 (1.5)	152 (1.3)
Nebraska	—	156 (1.3)
Nevada	140 (0.9)	137 (0.9)
New Mexico	141 (0.8)	140 (1.1)
New York ‡	146 (1.5) *,**	151 (1.6)
North Carolina	150 (1.5) *,**	157 (1.3)
North Dakota ‡	—	147 (1.2)
Ohio	—	160 (2.1)
Oklahoma	152 (1.3)	150 (1.2)
Oregon ‡	149 (1.5) *	155 (2.1)
Pennsylvania	—	154 (1.4)
Rhode Island	148 (0.7) *,**	151 (0.8)
South Carolina	140 (1.1) *,**	146 (1.1)
Tennessee ‡	148 (1.8)	148 (1.5)
Texas	154 (1.5)	152 (1.6)
Utah	143 (1.2)	143 (1.0)
Vermont	—	163 (1.2)
Virginia	153 (1.2)	157 (1.3)
Washington ‡	148 (1.5) *,**	155 (1.8)
West Virginia	144 (1.6)	144 (1.4)
Wisconsin ‡	153 (1.3)	—
Wyoming	146 (1.4) *,**	151 (0.9)
Other Jurisdictions		
American Samoa	—	95 (2.3)
District of Columbia	126 (1.2)	128 (0.8)
DDESS [2]	160 (2.6)	164 (1.5)
DoDDS [3]	156 (1.2) *,**	161 (0.8)
Guam	—	130 (1.4)
Virgin Islands	124 (3.8)	128 (1.2)

— Indicates that the jurisdiction did not participate or did not meet minimum participation guidelines for reporting.
‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
* Significantly different from 2002 when only one jurisdiction or the nation is being examined.
** Significantly different from 2002 when using a multiple-comparison procedure based on all jurisdictions that participated both years.
[1] National results that are presented for assessments prior to 2002 are based on the national sample, not on aggregated state assessment samples.
[2] Department of Defense Domestic Dependent Elementary and Secondary Schools. [3] Department of Defense Dependents Schools (Overseas).
NOTE: Comparative performance results may be affected by changes in exclusion rates for students with disabilities and limited English proficient students in the NAEP samples.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table A.12 **Percentage of students at or above proficient in writing and standard errors, by race/ethnicity, grade 8 public schools: By state, 1998 and 2002**

Grade 8	White		Black		Hispanic	
	1998	2002	1998	2002	1998	2002
Nation (Public) [1]	31 (1.0) *	37 (0.7)	7 (0.7) *	13 (0.6)	9 (1.2) *	15 (1.2)
Alabama	22 (1.4)	26 (2.0)	6 (1.5)	9 (1.5)	***(***)	***(***)
Arizona	28 (2.1)	27 (2.2)	6 (3.7)	13 (4.4)	7 (1.4)	9 (1.4)
Arkansas	16 (1.4) *	22 (1.8)	4 (1.1)	8 (2.2)	***(***)	12 (3.6)
California ‡	30 (2.5)	34 (2.8)	11 (3.3)	10 (3.1)	7 (1.3)	13 (2.6)
Colorado	32 (1.8)	—	10 (4.9)	—	9 (1.6)	—
Connecticut	52 (1.6)	55 (1.8)	14 (2.9)	15 (2.7)	13 (4.6)	17 (4.1)
Delaware	28 (2.2) *,**	43 (1.1)	9 (2.0) *,**	18 (1.3)	12 (4.5)	20 (4.2)
Florida	26 (2.3) *,**	41 (1.9)	7 (1.6) *,**	17 (2.4)	15 (3.2) *	26 (2.9)
Georgia	31 (2.3)	33 (2.0)	9 (1.3)	14 (2.1)	***(***)	7 (2.5)
Hawaii	20 (3.3)	21 (2.6)	***(***)	17 (6.2)	***(***)	***(***)
Idaho	—	30 (1.7)	—	***(***)	—	11 (2.8)
Indiana	—	29 (2.4)	—	7 (2.9)	—	***(***)
Kansas ‡	—	36 (1.8)	—	13 (4.5)	—	13 (4.5)
Kentucky	22 (1.8)	26 (1.6)	8 (2.9)	12 (2.6)	***(***)	***(***)
Louisiana	17 (1.3) *,**	26 (1.9)	4 (0.9) *,**	8 (1.1)	***(***)	***(***)
Maine	32 (1.7)	36 (1.5)	***(***)	***(***)	***(***)	***(***)
Maryland	31 (2.0) *,**	45 (2.2)	7 (2.0) *,**	17 (2.0)	12 (4.4)	24 (4.8)
Massachusetts	36 (2.1) *,**	49 (1.5)	9 (3.4)	18 (3.4)	6 (2.2)	10 (2.8)
Michigan	—	29 (1.8)	—	9 (2.5)	—	***(***)
Minnesota ‡	27 (2.2)	—	8 (3.5)	—	***(***)	—
Mississippi	17 (1.4)	20 (2.0)	4 (1.0)	6 (1.1)	***(***)	***(***)
Missouri	20 (1.6) *,**	29 (1.6)	4 (1.8) *	13 (3.5)	***(***)	***(***)
Montana ‡	26 (1.9)	32 (1.6)	***(***)	***(***)	***(***)	***(***)
Nebraska	—	35 (2.2)	—	10 (4.4)	—	11 (3.2)
Nevada	21 (1.2)	19 (1.2)	10 (3.8)	8 (2.2)	7 (1.8)	7 (1.4)
New Mexico	27 (2.1)	29 (2.0)	29 (8.6)	***(***)	11 (1.5)	13 (1.2)
New York ‡	29 (2.2) *,**	41 (2.5)	7 (2.4)	12 (2.4)	5 (1.8)	12 (2.8)
North Carolina	35 (2.2) *	43 (2.2)	11 (1.5) *,**	18 (1.9)	***(***)	16 (4.3)
North Dakota ‡	—	25 (1.6)	—	***(***)	—	***(***)
Ohio	—	42 (2.5)	—	14 (3.7)	—	***(***)
Oklahoma	29 (1.6)	31 (1.8)	7 (3.6)	13 (2.7)	13 (4.6)	13 (5.1)
Oregon ‡	28 (1.7) *	35 (2.4)	***(***)	***(***)	13 (4.5)	17 (3.9)
Pennsylvania	—	37 (1.8)	—	7 (1.5)	—	9 (2.6) !
Rhode Island	29 (1.5) *,**	35 (1.3)	10 (2.9)	10 (2.2)	5 (2.0)	9 (1.9)
South Carolina	22 (1.5) *,**	28 (1.9)	5 (1.3) *	9 (1.2)	***(***)	***(***)
Tennessee ‡	28 (2.1)	27 (2.0)	9 (2.2)	12 (2.8)	***(***)	***(***)
Texas	40 (2.1)	47 (2.7)	20 (3.9)	20 (3.1)	20 (2.3)	17 (2.0)
Utah	23 (1.2)	25 (1.1)	***(***)	***(***)	5 (2.6)	10 (2.5)
Vermont	—	42 (1.6)	—	***(***)	—	***(***)
Virginia	33 (1.7)	39 (2.2)	12 (1.7)	14 (1.7)	21 (6.0)	20 (6.0)
Washington ‡	27 (2.0) *,**	37 (2.4)	11 (4.7)	19 (5.2)	7 (2.6)	16 (3.0)
West Virginia	18 (1.7)	21 (1.4)	16 (5.9)	13 (5.3)	***(***)	***(***)
Wisconsin ‡	30 (1.8)	—	16 (3.8)	—	13 (5.4) !	—
Wyoming	24 (1.9)	30 (1.3)	***(***)	***(***)	14 (5.7)	12 (3.3)
Other Jurisdictions						
American Samoa	—	***(***)	—	***(***)	—	***(***)
District of Columbia	53 (10.3)	***(***)	9 (1.3)	8 (0.9)	10 (5.1)	11 (3.1)
DDESS [2]	47 (3.8)	51 (2.8)	27 (5.1)	27 (4.5)	32 (6.6)	38 (5.2)
DoDDS [3]	37 (2.7)	43 (1.8)	22 (4.0)	25 (2.8)	28 (5.2)	28 (4.3)
Guam	—	***(***)	—	***(***)	—	***(***)
Virgin Islands	***(***)	***(***)	8 (2.3)	4 (0.9)	7 (4.0)	2 (1.9)

See footnotes at end of table. ▶

Table A.12 **Percentage of students at or above proficient in writing and standard errors, by race/ethnicity, grade 8 public schools: By state, 1998 and 2002—Continued**

Grade 8	Asian/Pacific Islander		American Indian/Alaska Native		Other	
	1998	2002	1998	2002	1998	2002
Nation (Public) [1]	30 (5.8)	39 (2.8)	11 (4.6) !	17 (2.6)	20 (6.3) !	28 (5.0)
Alabama	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Arizona	***(***)	***(***)	12 (5.9)	8 (2.2) !	***(***)	***(***)
Arkansas	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
California ‡	35 (5.9)	36 (4.8)	***(***)	***(***)	***(***)	***(***)
Colorado	34 (8.9)	—	***(***)	—	***(***)	—
Connecticut	***(***)	55 (6.3)	***(***)	***(***)	***(***)	***(***)
Delaware	***(***)	63 (7.1)	***(***)	***(***)	***(***)	***(***)
Florida	***(***)	47 (8.1)	***(***)	***(***)	***(***)	***(***)
Georgia	***(***)	27 (6.2)	***(***)	***(***)	***(***)	***(***)
Hawaii	15 (1.2)	18 (1.0)	***(***)	***(***)	11 (2.7)	18 (2.7)
Idaho	—	***(***)	—	***(***)	—	***(***)
Indiana	—	***(***)	—	***(***)	—	***(***)
Kansas ‡	—	***(***)	—	***(***)	—	***(***)
Kentucky	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Louisiana	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Maine	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Maryland	40 (8.7)	55 (7.2)	***(***)	***(***)	***(***)	***(***)
Massachusetts	36 (8.1)	45 (6.2)	***(***)	***(***)	***(***)	***(***)
Michigan	—	***(***)	—	***(***)	—	***(***)
Minnesota ‡	11 (4.8)	—	***(***)	—	***(***)	—
Mississippi	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Missouri	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Montana ‡	***(***)	***(***)	14 (4.5) !	10 (1.6) !	***(***)	***(***)
Nebraska	—	***(***)	—	***(***)	—	***(***)
Nevada	18 (6.0)	28 (4.8)	***(***)	***(***)	***(***)	***(***)
New Mexico	***(***)	***(***)	12 (2.2)	9 (2.4)	***(***)	***(***)
New York ‡	27 (7.8) !	34 (7.5)	***(***)	***(***)	***(***)	***(***)
North Carolina	***(***)	***(***)	18 (6.4)	***(***)	***(***)	***(***)
North Dakota ‡	—	***(***)	—	7 (3.5) !	—	***(***)
Ohio	—	***(***)	—	***(***)	—	***(***)
Oklahoma	***(***)	***(***)	16 (4.0)	22 (2.6)	***(***)	***(***)
Oregon ‡	35 (6.2)	41 (7.5)	***(***)	***(***)	***(***)	***(***)
Pennsylvania	—	31 (10.4) !	—	***(***)	—	***(***)
Rhode Island	19 (6.2)	***(***)	***(***)	***(***)	***(***)	***(***)
South Carolina	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Tennessee ‡	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Texas	35 (7.0)	30 (9.2) !	***(***)	***(***)	***(***)	***(***)
Utah	16 (5.6)	17 (5.8)	***(***)	***(***)	***(***)	***(***)
Vermont	—	***(***)	—	***(***)	—	***(***)
Virginia	40 (7.7)	46 (7.2)	***(***)	***(***)	***(***)	***(***)
Washington ‡	27 (5.9)	35 (4.6)	***(***)	***(***)	***(***)	***(***)
West Virginia	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
Wisconsin ‡	***(***)	—	***(***)	—	***(***)	—
Wyoming	***(***)	***(***)	8 (5.2) !	13 (4.9)	***(***)	***(***)
Other Jurisdictions						
American Samoa	—	3 (1.3)	—	***(***)	—	***(***)
District of Columbia	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)
DDESS [2]	***(***)	***(***)	***(***)	***(***)	***(***)	45 (8.4)
DoDDS [3]	30 (7.1)	35 (6.3)	***(***)	***(***)	29 (3.0)	38 (3.0)
Guam	—	13 (1.4)	—	***(***)	—	***(***)
Virgin Islands	***(***)	***(***)	***(***)	***(***)	***(***)	***(***)

— Indicates that the jurisdiction did not participate or did not meet minimum participation guidelines for reporting.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
* Significantly different from 2002 when only one jurisdiction or the nation is being examined.
** Significantly different from 2002 when using a multiple-comparison procedure based on all jurisdictions that participated both years.
***(***) Sample size is insufficient to permit a reliable estimate.
[1] National results that are presented for assessments prior to 2002 are based on the national sample, not on aggregated state assessment samples.
[2] Department of Defense Domestic Dependent Elementary and Secondary Schools. [3] Department of Defense Dependents Schools (Overseas).
Comparative performance results may be affected by changes in exclusion rates for students with disabilities and limited English proficient students in the NAEP samples.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Analyzing Group Differences in Averages and Percentages

Statistical tests determine whether the evidence, based on the data from the groups in the sample, is strong enough to conclude that the averages or percentages are actually different for those groups in the population. If the evidence is strong (i.e., the difference is statistically significant), the report describes the group averages or percentages as being different (e.g., one group performed higher or lower than another group), regardless of whether the sample averages or percentages appear to be approximately the same. The reader is cautioned to rely on the results of the statistical tests rather than on the apparent magnitude of the difference between sample averages or percentages when determining whether the sample differences are likely to represent actual differences among the groups in the population.

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

Standard Error of the Difference =

$$SE_{A-B} =$$

The standard error of the difference can be used, just like the standard error for an individual group average or percentage, to help determine whether differences among groups in the population are real. The difference between the averages or percentages of the two groups plus or minus 1.96 standard errors of the difference represents an approximately 95 percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between the groups in the population. If the interval does not contain zero, the difference between the groups is statistically significant at the 0.05 level.

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

Group	Average Scale Score	Standard Error
A	137	0.9
B	135	1.1

The difference between the estimates of the average scale scores of groups A and B is two points (137−135). The estimated standard error of this difference is

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

Thus, an approximately 95 percent confidence interval for this difference is plus or minus 1.96 standard errors of the difference.

$$2 \pm 1.96 \times 1.4$$

$$2 \pm 2.7$$

$$(-0.7, 4.7)$$

The value zero is within the confidence interval; therefore, there is insufficient evidence to claim that group A outperformed group B.

The procedure above is appropriate to use when it is reasonable to assume that the groups being compared have been independently sampled for the assessment. Such an assumption is clearly warranted when comparing results across assessment years (e.g., comparing the 1998 and 2002 results for a particular state or subgroup) or when comparing state results with each other. This is the approach used for NAEP reports when comparisons involving independent groups are made. The assumption of independence is violated to some degree when comparing group results for the nation or a particular state (e.g., comparing national 2002 results for males and females), since these samples of students have been drawn from the same schools. When the groups being compared do not share students (as is the case, for example, comparing males and females) the impact of this violation of the independence assumption on the outcome of the statistical tests is assumed to be small, and NAEP, by convention, has, for computational convenience, routinely applied the procedures described above to those cases as well.

When making comparisons of results for groups that share a considerable proportion of students in common, it is not appropriate to ignore such dependencies. In such cases, NAEP has used procedures appropriate to comparing dependent groups. When the dependence in group results is due to the overlap in samples (e.g., when a subgroup is being compared to a total group), a simple modification of the usual standard error of the difference formula can be used. The formula for such cases is:[15]

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

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

[15] This is a special form of the common formula for standard error of dependent samples. The standard formula can be found, for example, in Klish, L. (1995). *Survey Sampling.* New York: John Wiley and Sons, Inc.

Conducting Multiple Tests

The procedures in the previous 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. However, there are times when many different groups are being compared (i.e., multiple sets of confidence intervals are being analyzed). In 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. To hold the significance level for the set of comparisons at a particular level (e.g., 0.05), adjustments (called "multiple comparison procedures")[16] must be made to the methods described in the previous section. One such procedure, the Benjamini-Hochberg False Discovery Rate (FDR) procedure was used to control the certainty level.[17]

Unlike the other multiple comparison procedures that control the familywise error rate (i.e., the probability of making even one false rejection in the set of comparisons), the FDR procedure controls the expected proportion of falsely rejected hypotheses. Furthermore, the FDR procedure used in NAEP is considered appropriately less conservative than familywise procedures for large families of comparisons.[18] Therefore, the FDR procedure is more suitable for multiple comparisons in NAEP than other procedures. A detailed description of the FDR procedure will appear in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

To illustrate how the FDR procedure is used, consider the comparisons of current and previous years' average writing scale scores for the five groups presented in table A.13. Note that the difference in average scale scores and the estimated standard error of the difference are calculated in a way comparable to that of the example in the previous section. The test statistic shown is the difference in average scale scores divided by the estimated standard error of the difference. (Rounding of the data occurs after the test is done.)

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

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

18 Williams, V. S. L., Jones, L. V., and Tukey, J. W. (1999). Controlling Error in Multiple Comparisons with Examples From State-to-State Differences in Educational Achievement. *Journal of Educational and Behavioral Statistics*, *24*(1), 42–69.

Table A.13 Example of FDR comparisons of average scale scores for different groups of students

	Previous year		Current year		Previous year and current year			
	Average scale score	Standard error	Average scale score	Standard error	Difference in averages	Standard error of difference	Test statistic	Percent confidence[1]
Group 1	224	1.3	226	1.0	2.08	1.62	1.29	20
Group 2	187	1.7	193	1.7	6.31	2.36	2.68	1
Group 3	191	2.6	197	1.7	6.63	3.08	2.15	4
Group 4	229	4.4	232	4.6	3.24	6.35	0.51	62
Group 5	201	3.4	196	4.7	-5.51	5.81	-0.95	35

[1] The percent confidence is 2(1-F(x)) where F(x) is the cumulative distribution of the t-distribution with the degrees of freedom adjusted to reflect the complexities of the sample design.
FDR: False Discovery Rate.

The difference in average scale scores and its estimated standard error can be used to find an approximately 95 percent confidence interval as in the example in the previous section or they can be used to identify a confidence percentage. In the example in the previous section, because an approximately 95 percent confidence interval was desired, the number 1.96 was used to multiply the estimated standard error of the difference to create the approximate confidence interval. In the current example, the confidence interval for the test statistics is identified from statistical tables. Instead of checking to see if zero is within the 95 percent confidence interval about the mean, the significance level from the statistical tables can be directly compared to 100 − 95 = 5 percent.

If the comparison of average scale scores across two years was made for only one of the five groups, there would be a significant difference between the average scale scores for the two years if the significance level were less than 5 percent. However, because we are interested in the difference in average scale scores across the two years for all five of the groups, comparing each of the significance levels to 5 percent is not adequate. Groups of students defined by shared characteristics, such as racial/ethnic groups, are treated as sets or families when making comparisons. However, comparisons of average scale scores for each pair of years were treated separately, so the steps described in this example would be replicated for the comparison of other current and previous year average scale scores.

Using the FDR procedure to take into account that all comparisons are of interest to us, the percents of confidence in the example are ordered from largest to smallest: 62, 35, 20, 4, and 1. In the FDR procedure, 62 percent confidence for the group 4 comparison would be compared to 5 percent, 35 percent for the group 5 comparison would be compared to $0.05 \times (5-1)/5 = 0.04 = 4$ percent,[19] 20 percent for the group 1 comparison would be compared to $0.05 \times (5-2)/5 = 0.03 = 3$ percent, 4 percent for the group 3 comparison would be compared to $0.05 \times (5-3)/5 = 0.02 = 2$ percent, and 1 percent for the group 2 comparison (actually slightly smaller than 1 prior to rounding) would be compared to $0.05 \times$

[19] The level of confidence times the number of comparisons minus one divided by the number of comparisons is $0.05 \times (5-1)/5 = 0.04 = 4$ percent.

$(5-4)/5 = 0.01 = 1$ percent. The procedure stops with the first contrast found to be significant. The last of these comparisons is the only one for which the percent confidence is smaller than the FDR procedure value. The difference in the current year and previous years' average scale scores for the group 2 students is significant; for all of the other groups, average scale scores for current and previous year are not significantly different from one another. In practice, a very small number of counterintuitive results occur when the FDR procedures are used to examine between-year differences in subgroup results by jurisdiction. In those cases, results were not included in this report.

NAEP Reporting Groups

Results are provided for groups of students defined by shared characteristics—gender, race or ethnicity, school's type of location, Title I participation, eligibility for free/reduced-price school lunch, and type of school. Based on participation rate criteria, results are reported for subpopulations only when sufficient numbers of students and adequate school representation are present. The minimum requirement is at least 62 students in a particular subgroup from at least five primary sampling units (PSUs).[20] However, the data for all students, regardless of whether their subgroup was reported separately, were included in computing overall results. Definitions of the subpopulations are presented below.

Gender

Results are reported separately for males and females.

Race/Ethnicity

In all NAEP assessments, data about student race/ethnicity is collected from two sources: school records and student self-reports. Previously, NAEP has used student self-reported race as the primary race/ethnicity reporting variable. In 2002, it was decided to change the student race/ethnicity variable highlighted in NAEP reports. Starting in 2002, school-recorded race will become the race/ethnicity variable presented in NAEP reports. Information based on student self-reported race/ethnicity will continue to be available on the NAEP Data Tool (*http://nces.ed.gov/nationsreportcard/naepdata*).

In order to allow comparisons across years, both the 1998 and 2002 writing assessment results presented in this report are based on school-reported information for six mutually exclusive racial/ethnic categories: White, Black, Hispanic, Asian/Pacific Islander, American Indian (including Alaska Native), and Other. Students who were identified with more than one of the first five categories or had a background other than the ones listed were categorized as Other. Information about the percentage of students identified as Other is presented in tables B.12 and B.13 in appendix B.

Type of Location

Results from the 2002 assessment are reported for students attending schools in three mutually exclusive location types:

Central city: This category includes central cities of all Consolidated Metropolitan Statistical Area (CMSA) or Metropolitan Statistical Area (MSA) as defined by the

[20] For the NAEP national assessments prior to 2002, a PSU is a selected geographic region (a county, group of counties, or metropolitan statistical area). In 2002, the first-stage sampling units are schools (public and nonpublic) in the selection of the combined sample. Further details about the procedure for determining minimum sample size will appear in the technical documentation section of the NAEP web site at *http://nces.ed.gov/nationsreportcard*.

Office of Management and Budget. Central city is a geographical term and is not synonymous with "inner city."

Urban fringe/large town: The urban fringe category includes any incorporated place, census designated place, or non-place territory within a CMSA or MSA of a large or mid-sized city and defined as urban by the U.S. Census Bureau, but which do not qualify as central city. A large town is defined as a place outside a CMSA or MSA 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. Census Bureau. A small town is defined as a place outside a CMSA or MSA with a population of less than 25,000, but greater than or equal to 2,500.

Results for each type of location are not compared across years. This is due to new methods used by NCES to identify the type of location assigned to each school in the Common Core of Data (CCD). The new methods were put into place by NCES in order to improve the quality of the assignments, and they take into account more information about the exact physical location of the school. The variable was revised in NAEP beginning with the 2000 assessments.

Title I Participation

Based on available school records, students were classified either as currently participating in a Title I program, receiving Title I services, or as not receiving such services. The classification applies only to the school year when the assessment was administered (i.e., the 2001–02 school year) and is not based on participation in previous years. If the school does not offer any Title I programs or services, all students in that school would be classified as not participating.

Eligibility for Free/Reduced-Price School Lunch

As part of the Department of Agriculture's National School Lunch Program, schools can receive cash subsidies and donated commodities in turn for offering free or reduced-price lunches to eligible children. Based on available school records, students were classified as either currently eligible for the free/reduced-price school lunch or not eligible. Eligibility for free and reduced-price lunches is determined by students' family income in relation to the federally established poverty level. Free lunch qualification is set at 130 percent of the poverty level, and reduced-price lunch qualification is set at 170 percent of the poverty level. The classification applies only to the school year when the assessment was administered (i.e., the 2001–02 school year) and is not based on eligibility in previous years. If school records were not available, the student was classified as "Information not available." If the school did not participate in the program, all students in that school were classified as "Information not available."

Type of School

Results are reported by the type of school that the student attends—public or nonpublic. Nonpublic schools include Catholic and other private schools. Because they are funded by federal authorities, not state/local governments, Bureau of Indian Affairs (BIA) schools and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) are not included in either the public or nonpublic categories; they are included in the overall national results.

Grade 12 Participation Rates

NAEP has been described as a "low-stakes" assessment. That is, students receive no individual scores, and their NAEP performance has no affect on their grades, promotions, or graduation. There has been continued concern that this lack of consequences affects participation rates of students and schools, as well as the motivation of students to perform well on NAEP. Of particular concern has been the performance of twelfth-graders, who typically have lower student participation rates than fourth- and eighth-graders and who are more likely to omit responses compared to their younger cohorts.

In NAEP, there has been a consistent pattern of lower participation rates for older students. In the 2002 NAEP assessments, for example, the student participation rates were 94 percent and 92 percent at grades 4 and 8, respectively. At grade 12, however, the participation rate was 74 percent. School participation rates (the percentage of sampled schools that participated in the assessment) have also typically decreased with grade level. In the 2002 assessments, the school participation rate was 85 percent for the fourth grade, 83 percent for the eighth grade, and 75 percent for the twelfth grade.

The effect of participation rates on student performance, however, is unclear. Students may choose not to participate in NAEP for many reasons such as desire to attend regular classes and not miss important instruction or conflict with other school-based activities. Similarly, there are a variety of reasons for which various schools do not participate. The sampling weights and nonresponse adjustments, described earlier in this document, provide an approximate statistical adjustment for nonparticipation. However, the effect of some school and student nonparticipation may have some undetermined effect on results.

More research is needed to delineate the factors that contribute to nonparticipation and lack of motivation. To that end, NCES is currently investigating how various types of incentives can be effectively used to increase participation in NAEP.

Cautions in Interpretations

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

A caution is also warranted for some small population group estimates. At times in this report, smaller population groups show very large increases or decreases across years in average scores. However, it is often necessary to interpret such score gains with extreme caution. For one thing, the effects of exclusion-rate changes for small subgroups may be more marked for small groups than they are for the whole population. Also, the standard errors are often quite large around the score estimates for small groups, which in turn means the standard error around the gain is also large.

B Appendix B
Subgroup Percentages

Table B.1 **Weighted percentage of students, by gender, grades 4, 8, and 12: 1998 and 2002**

		1998	2002
Grade 4			
	Male	51	51
	Female	49	49
Grade 8			
	Male	51	50
	Female	49	50
Grade 12			
	Male	48	49
	Female	52	51

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table B.2 **Weighted percentage of students, by race/ethnicity, grades 4, 8, and 12: 1998 and 2002**

	1998	2002
Grade 4		
White	71	61
Black	16	17
Hispanic	9	16
Asian/Pacific Islander	3	4
American Indian/Alaska Native	1	1
Other	1	1
Grade 8		
White	70	65
Black	15	15
Hispanic	11	14
Asian/Pacific Islander	3	4
American Indian/Alaska Native	1	1
Other	#	1
Grade 12		
White	72	70
Black	14	13
Hispanic	10	10
Asian/Pacific Islander	4	5
American Indian/Alaska Native	#	#
Other	#	1

Percentage rounds to zero.
NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table B.3 **Weighted percentage of students, by eligibility for free/reduced-price school lunch, grades 4, 8, and 12: 1998 and 2002**

	1998	2002
Grade 4		
Eligible	34	40
Not eligible	54	47
Information not available	13	13
Grade 8		
Eligible	27	31
Not eligible	55	53
Information not available	17	15
Grade 12		
Eligible	14	19
Not eligible	66	64
Information not available	20	17

NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table B.4 Weighted percentage of students, by eligibility for free/reduced-price school lunch and race/ethnicity, grades 4, 8, and 12: 2002

	Eligible	Not eligible	Information not available
Grade 4			
White	24	62	14
Black	69	23	8
Hispanic	68	19	13
Asian/Pacific Islander	33	47	20
American Indian/Alaska Native	60	31	9
Grade 8			
White	20	65	16
Black	58	30	12
Hispanic	58	28	14
Asian/Pacific Islander	31	45	24
American Indian/Alaska Native	51	37	12
Grade 12			
White	11	71	18
Black	44	44	12
Hispanic	43	41	17
Asian/Pacific Islander	24	59	16
American Indian/Alaska Native	***	***	***

*** Quality control activities and special analysis raised concerns about the accuracy and precision of grade 12 American Indian data in 2002. As a result, they are omitted from this report.
NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.5 **Weighted percentage of students, by school participation in Title I, grades 4, 8, and 12: 2002**

	2002
Grade 4	
Participated	33
Did not participate	67
Grade 8	
Participated	19
Did not participate	81
Grade 12	
Participated	10
Did not participate	90

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.6 **Weighted percentage of students, by student-reported parents' highest level of education, grades 8 and 12: 2002**

	2002
Grade 8	
Less than high school	7
Graduated high school	17
Some education after high school	19
Graduated college	48
Unknown	9
Grade 12	
Less than high school	7
Graduated high school	18
Some education after high school	25
Graduated college	47
Unknown	3

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.7 **Weighted percentage of students, by type of school, grades 4, 8, and 12: 1998 and 2002**

	1998	2002
Grade 4		
Public	88	90
Nonpublic	12	10
Nonpublic: Catholic	7	5
Nonpublic: Other	4	4
Grade 8		
Public	89	91
Nonpublic	11	9
Nonpublic: Catholic	7	5
Nonpublic: Other	5	4
Grade 12		
Public	88	91
Nonpublic	12	9
Nonpublic: Catholic	8	5
Nonpublic: Other	3	5

NOTE: Percentages may not add to 100, or to the exact nonpublic percentages, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table B.8 **Weighted percentage of students, by student-reported parents' highest level of education and type of school, grades 8 and 12: 2002**

	Less than high school	Graduated high school	Some education after high school	Graduated college	Unknown
Grade 8					
Public	7	18	20	46	10
Nonpublic	2	9	15	69	5
Grade 12					
Public	7	19	25	45	3
Nonpublic	2	11	20	66	2

NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.9 **Weighted percentage of students, by type of location, grades 4, 8, and 12: 2002**

	2002
Grade 4	
Central city	30
Urban fringe/large town	42
Rural/small town	28
Grade 8	
Central city	29
Urban fringe/large town	42
Rural/small town	29
Grade 12	
Central city	29
Urban fringe/large town	40
Rural/small town	31

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.10 Weighted percentage of students, by gender, grade 4: By state, 2002

Grade 4	Male	Female
Nation (Public)	51	49
Alabama	51	49
Arizona	51	49
Arkansas	49	51
California ‡	52	48
Connecticut	52	48
Delaware	50	50
Florida	51	49
Georgia	51	49
Hawaii	51	49
Idaho	53	47
Indiana	50	50
Iowa ‡	52	48
Kansas ‡	49	51
Kentucky	50	50
Louisiana	52	48
Maine	51	49
Maryland	48	52
Massachusetts	52	48
Michigan	51	49
Minnesota ‡	51	49
Mississippi	50	50
Missouri	50	50
Montana ‡	51	49
Nebraska	50	50
Nevada	49	51
New Mexico	53	47
New York ‡	51	49
North Carolina	50	50
North Dakota ‡	50	50
Ohio	50	50
Oklahoma	51	49
Oregon	50	50
Pennsylvania	51	49
Rhode Island	52	48
South Carolina	51	49
Tennessee ‡	50	50
Texas	51	49
Utah	52	48
Vermont	50	50
Virginia	50	50
Washington ‡	54	46
West Virginia	49	51
Wyoming	51	49
Other Jurisdictions		
District of Columbia	49	51
DDESS [1]	50	50
DoDDS [2]	50	50
Guam	52	48
Virgin Islands	49	51

‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.11 Weighted percentage of students, by gender, grade 8: By state, 1998 and 2002

Grade 8	Male		Female	
	1998	2002	1998	2002
Nation (Public)	51	50	49	50
Alabama	49	50	51	50
Arizona	51	50	49	50
Arkansas	50	53	50	47
California ‡	48	52	52	48
Colorado	51	—	49	—
Connecticut	50	51	50	49
Delaware	51	51	49	49
Florida	49	50	51	50
Georgia	52	51	48	49
Hawaii	53	52	47	48
Idaho	—	52	—	48
Indiana	—	50	—	50
Kansas ‡	—	51	—	49
Kentucky	50	49	50	51
Louisiana	47	51	53	49
Maine	49	49	51	51
Maryland	50	48	50	52
Massachusetts	51	53	49	47
Michigan	—	52	—	48
Minnesota ‡	51	—	49	—
Mississippi	49	49	51	51
Missouri	51	50	49	50
Montana ‡	50	53	50	47
Nebraska	—	51	—	49
Nevada	50	52	50	48
New Mexico	52	51	48	49
New York ‡	51	52	49	48
North Carolina	51	50	49	50
North Dakota ‡	—	52	—	48
Ohio	—	50	—	50
Oklahoma	52	50	48	50
Oregon ‡	51	53	49	47
Pennsylvania	—	51	—	49
Rhode Island	51	52	49	48
South Carolina	51	50	49	50
Tennessee ‡	48	51	52	49
Texas	49	50	51	50
Utah	49	51	51	49
Vermont	—	52	—	48
Virginia	52	51	48	49
Washington ‡	49	52	51	48
West Virginia	52	51	48	49
Wisconsin ‡	51	—	49	—
Wyoming	52	51	48	49
Other Jurisdictions				
American Samoa	—	50	—	50
District of Columbia	48	49	52	51
DDESS [1]	51	47	49	53
DoDDS [2]	49	50	51	50
Guam	—	51	—	49
Virgin Islands	44	47	56	53

— Indicates that the jurisdiction did not participate or did not meet minimum participation guidelines for reporting.
‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table B.12 Weighted percentage of students, by race/ethnicity, grade 4: By state, 2002

Grade 4	White	Black	Hispanic	Asian/ Pacific Islander	American Indian/ Alaska Native	Other
Nation (Public)	60	18	17	4	1	1
Alabama	61	36	1	1	1	#
Arizona	50	5	35	2	7	#
Arkansas	69	25	5	1	#	#
California ‡	35	7	46	10	1	1
Connecticut	72	13	11	3	#	#
Delaware	58	33	6	2	#	#
Florida	51	24	22	2	#	1
Georgia	53	38	5	3	#	1
Hawaii	17	3	3	63	#	13
Idaho	85	1	11	1	2	#
Indiana	80	13	4	1	1	1
Iowa ‡	86	6	4	2	#	1
Kansas ‡	78	8	10	2	1	#
Kentucky	86	12	1	1	#	1
Louisiana	46	51	2	1	1	#
Maine	96	2	1	1	#	#
Maryland	52	37	5	4	1	#
Massachusetts	78	8	8	5	#	1
Michigan	72	20	4	2	2	1
Minnesota ‡	82	6	4	4	4	1
Mississippi	47	52	1	1	#	#
Missouri	79	17	2	1	#	#
Montana ‡	86	1	2	1	10	#
Nebraska	82	6	8	1	3	#
Nevada	53	11	28	6	2	#
New Mexico	35	2	47	2	13	1
New York ‡	54	19	21	6	#	1
North Carolina	58	31	6	2	2	2
North Dakota ‡	88	1	1	1	8	#
Ohio	76	20	2	1	#	1
Oklahoma	59	11	8	1	18	2
Oregon	78	3	11	5	1	2
Pennsylvania	77	17	4	2	#	#
Rhode Island	73	9	13	3	1	#
South Carolina	55	42	2	1	#	#
Tennessee ‡	73	23	2	1	#	#
Texas	36	18	41	3	1	#
Utah	85	1	8	3	1	#
Vermont	96	1	1	1	#	1
Virginia	64	25	5	4	1	1
Washington ‡	77	7	6	7	3	#
West Virginia	95	4	#	1	#	#
Wyoming	86	1	8	1	4	1
Other Jurisdictions						
District of Columbia	4	87	7	1	#	#
DDESS [1]	40	27	12	3	1	16
DoDDS [2]	47	15	7	7	1	22
Guam	1	#	#	98	#	#
Virgin Islands	1	86	12	#	#	1

Percentage rounds to zero.
‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.13 Weighted percentage of students, by race/ethnicity, grade 8: By state, 1998 and 2002

Grade 8	White		Black		Hispanic		Asian/ Pacific Islander		American Indian/ Alaska Native		Other	
	1998	2002	1998	2002	1998	2002	1998	2002	1998	2002	1998	2002
Nation (Public)	69	64	16	15	11	14	3	4	1	1	#	1
Alabama	67	62	31	36	1	1	1	1	#	#	#	#
Arizona	60	57	4	5	26	30	2	2	7	6	#	#
Arkansas	74	73	23	23	2	3	1	1	#	1	#	#
California ‡	42	37	8	7	39	42	10	13	1	1	1	1
Colorado	75	—	5	—	17	—	3	—	1	—	#	—
Connecticut	78	70	11	14	9	12	2	3	#	1	1	1
Delaware	67	64	27	29	4	5	2	2	#	#	#	#
Florida	56	55	28	23	14	18	2	2	#	#	#	1
Georgia	58	54	36	37	2	5	2	3	#	#	1	1
Hawaii	17	16	2	2	2	2	67	68	#	#	12	12
Idaho	—	88	—	1	—	9	—	1	—	1	—	#
Indiana	—	86	—	9	—	2	—	1	—	#	—	1
Kansas ‡	—	80	—	8	—	7	—	2	—	1	—	#
Kentucky	89	90	10	8	#	1	1	1	#	#	#	#
Louisiana	58	53	40	43	1	1	1	1	#	1	#	#
Maine	97	97	1	1	#	1	1	1	#	#	#	#
Maryland	59	55	34	34	3	5	4	5	#	#	#	#
Massachusetts	81	75	6	9	9	10	4	5	#	#	#	1
Michigan	—	77	—	18	—	2	—	2	—	#	—	#
Minnesota ‡	85	—	5	—	2	—	5	—	3	—	#	—
Mississippi	51	52	48	47	#	#	1	#	#	#	#	#
Missouri	84	81	14	16	1	1	1	1	#	#	#	#
Montana ‡	92	84	#	1	1	2	1	1	5	12	#	#
Nebraska	—	84	—	6	—	7	—	1	—	1	—	#
Nevada	65	60	9	10	19	22	5	7	2	1	#	#
New Mexico	40	36	3	2	46	47	1	1	9	13	1	#
New York ‡	60	55	19	21	15	17	5	6	#	#	1	#
North Carolina	64	63	28	30	2	4	2	2	3	#	#	1
North Dakota ‡	—	92	—	1	—	2	—	1	—	4	—	#
Ohio	—	80	—	15	—	2	—	1	—	#	—	2
Oklahoma	74	62	7	11	4	6	2	1	12	18	1	1
Oregon ‡	85	82	2	2	6	8	4	5	2	2	1	1
Pennsylvania	—	81	—	13	—	4	—	3	—	#	—	#
Rhode Island	81	75	7	9	8	13	3	2	#	#	1	#
South Carolina	58	56	40	42	1	1	1	1	#	#	#	#
Tennessee ‡	77	77	21	20	1	2	1	1	#	#	#	#
Texas	50	44	13	12	32	40	3	3	1	1	#	#
Utah	89	86	1	1	6	8	3	3	1	2	#	#
Vermont	—	96	—	1	—	#	—	1	—	1	—	#
Virginia	68	66	26	24	3	4	3	4	#	#	#	#
Washington ‡	81	79	4	4	7	7	6	8	2	2	#	#
West Virginia	95	95	4	4	#	#	#	#	#	#	#	#
Wisconsin ‡	84	—	8	—	4	—	3	—	1	—	#	—
Wyoming	90	88	1	2	5	7	1	1	2	3	#	#
Other Jurisdictions												
American Samoa	—	#	—	#	—	#	—	100	—	#	—	#
District of Columbia	4	3	89	87	5	8	1	2	#	#	#	#
DDESS [1]	42	38	27	23	22	20	2	6	1	1	7	13
DoDDS [2]	49	48	19	15	7	7	8	9	1	1	17	19
Guam	—	2	—	#	—	#	—	96	—	#	—	2
Virgin Islands	#	#	87	85	11	12	#	#	#	#	2	2

— Indicates that the jurisdiction did not participate or did not meet minimum participation guidelines for reporting.
Percentage rounds to zero.
‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Table B.14 Weighted percentage of students, by eligibility for free/reduced-price school lunch, grade 4: By state, 2002

Grade 4	Eligible	Not eligible	Information not available
Nation (Public)	43	49	7
Alabama	53	34	13
Arizona	46	36	18
Arkansas	56	40	3
California ‡	46	37	17
Connecticut	27	66	6
Delaware	38	60	2
Florida	55	43	2
Georgia	47	50	3
Hawaii	47	52	1
Idaho	45	47	9
Indiana	33	60	7
Iowa ‡	30	70	#
Kansas ‡	43	56	#
Kentucky	48	50	2
Louisiana	61	31	8
Maine	31	63	6
Maryland	39	58	3
Massachusetts	27	67	6
Michigan	38	57	5
Minnesota ‡	29	58	14
Mississippi	65	25	10
Missouri	41	56	3
Montana ‡	38	57	5
Nebraska	40	56	4
Nevada	38	56	6
New Mexico	56	29	15
New York ‡	44	49	7
North Carolina	49	47	4
North Dakota ‡	31	66	2
Ohio	32	61	7
Oklahoma	55	42	3
Oregon	38	48	13
Pennsylvania	34	63	3
Rhode Island	33	54	13
South Carolina	54	40	5
Tennessee ‡	45	50	5
Texas	58	37	5
Utah	32	63	5
Vermont	27	69	5
Virginia	33	65	3
Washington ‡	32	59	9
West Virginia	52	45	3
Wyoming	41	56	4
Other Jurisdictions			
District of Columbia	78	21	1
DDESS [1]	32	35	33
DoDDS [2]	8	25	66
Guam	61	39	#
Virgin Islands	99	#	1

Percentage rounds to zero.
‡ Indicates that the jurisdiction did not meet one or more of the guidelines for school participation in 2002.
[1] Department of Defense Domestic Dependent Elementary and Secondary Schools.
[2] Department of Defense Dependents Schools (Overseas).
NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessment.

Table B.15 Weighted percentage of students, by eligibility for free/reduced-price school lunch, grade 8: By state, 1998 and 2002

Grade 8	Eligible		Not eligible		Information not available	
	1998	2002	1998	2002	1998	2002
Nation (Public)	30	34	58	56	12	10
Alabama	39	42	59	42	2	16
Arizona	33	34	52	53	15	14
Arkansas	35	44	60	54	5	2
California [1‡]	39	36	45	46	17	18
Colorado	24	—	65	—	11	—
Connecticut	18	30	68	62	13	8
Delaware	27	32	63	68	11	1
Florida	40	43	50	52	10	5
Georgia	35	40	53	55	12	5
Hawaii	37	40	59	59	4	1
Idaho	—	32	—	60	—	8
Indiana	—	25	—	69	—	6
Kansas ‡	—	30	—	67	—	3
Kentucky	39	40	57	57	4	3
Louisiana	48	50	43	36	9	14
Maine	26	24	66	69	7	7
Maryland	28	26	69	71	3	2
Massachusetts	23	29	73	69	5	2
Michigan	—	34	—	60	—	7
Minnesota ‡	23	—	70	—	7	—
Mississippi	51	58	42	36	7	6
Missouri	28	30	69	65	3	6
Montana ‡	24	31	67	67	9	2
Nebraska	—	35	—	63	—	2
Nevada	26	28	65	62	9	9
New Mexico	43	51	42	29	15	20
New York ‡	37	37	46	56	17	8
North Carolina	32	38	61	53	7	9
North Dakota ‡	—	25	—	74	—	2
Ohio	—	24	—	65	—	11
Oklahoma	34	45	57	50	9	5
Oregon ‡	26	26	69	63	5	11
Pennsylvania	—	30	—	69	—	#
Rhode Island	27	24	71	60	1	16
South Carolina	41	45	55	51	4	4
Tennessee ‡	33	38	65	52	2	10
Texas	38	45	59	48	3	7
Utah	22	24	67	66	11	9
Vermont	—	21	—	78	—	1
Virginia	23	26	70	70	7	3
Washington ‡	23	22	67	56	10	22
West Virginia	39	44	57	55	3	1
Wisconsin ‡	21	—	71	—	8	—
Wyoming	24	32	74	65	2	3
Other Jurisdictions						
American Samoa	—	100	—	#	—	#
District of Columbia	61	67	21	32	17	1
DDESS [2]	33	25	65	54	2	21
DoDDS [3]	5	6	22	23	73	71
Guam	—	30	—	69	—	1
Virgin Islands	80	99	#	#	20	1

— Indicates that the jurisdiction did not participate or did not meet minimum participation guidelines for reporting.
Percentage rounds to zero.
‡ Indicates that the jurisdiction or national aggregate did not meet one or more of the guidelines for school participation in 2002.
1 Percentages by student's eligibility for free/reduced-price lunch in California do not include Los Angeles.
2 Department of Defense Domestic Dependent Elementary and Secondary Schools.
3 Department of Defense Dependents Schools (Overseas).
NOTE: Percentages may not add to 100, due to rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 and 2002 Writing Assessments.

Appendix C
State-Level Contextual Variables

To help place state results from the NAEP 2002 writing assessment into context, this appendix presents selected state-level data from the *Digest of Education Statistics 2001*.

Table C.1 Population and public-school enrollment, from non-NAEP sources: By state, April 2000 and fall 1999

	Estimated resident populations: April 1, 2000		Enrollment in public elementary and secondary schools: Fall 1999		
	Total (in thousands)	5- to 17-year-olds (in thousands)	Total	Kindergarten through grade 8[1]	Grades 9–12
Nation	281,422	53,118	46,857,321	33,488,158	13,369,163
Alabama	4,447	827	740,732	538,687	202,045
Alaska	627	143	134,391	95,601	38,790
Arizona	5,131	985	852,612	623,561	229,051
Arkansas	2,673	499	451,034	317,714	133,320
California	33,872	6,763	6,038,589	4,336,687	1,701,902
Colorado	4,301	803	708,109	506,568	201,541
Connecticut	3,406	618	553,993	403,913	150,080
Delaware	784	143	112,836	80,274	32,562
District of Columbia	572	82	77,194	59,917	17,277
Florida	15,982	2,701	2,381,396	1,725,493	655,903
Georgia	8,186	1,574	1,422,762	1,044,030	378,732
Hawaii	1,212	218	185,860	133,250	52,610
Idaho	1,294	271	245,331	168,822	76,509
Illinois	12,419	2,369	2,027,600	1,462,234	565,366
Indiana	6,080	1,151	988,702	699,221	289,481
Iowa	2,926	545	497,301	335,919	161,382
Kansas	2,688	524	472,188	325,818	146,370
Kentucky	4,042	729	648,180	458,607	189,573
Louisiana	4,469	902	756,579	548,019	208,560
Maine	1,275	231	209,253	148,774	60,479
Maryland	5,296	1,003	846,582	607,125	239,457
Massachusetts	6,349	1,103	971,425	706,251	265,174
Michigan	9,938	1,924	1,725,617	1,244,586	481,031
Minnesota	4,919	957	854,034	580,363	273,671
Mississippi	2,845	571	500,716	365,357	135,359
Missouri	5,595	1,058	914,110	648,758	265,352
Montana	902	175	157,556	107,490	50,066
Nebraska	1,711	333	288,261	197,014	91,247
Nevada	1,998	366	325,610	239,625	85,985
New Hampshire	1,236	234	206,783	146,854	59,929
New Jersey	8,414	1,524	1,289,256	953,766	335,490
New Mexico	1,819	378	324,495	228,592	95,903
New York	18,976	3,451	2,887,776	2,033,748	854,028
North Carolina	8,049	1,425	1,275,925	934,725	341,200
North Dakota	642	121	112,751	74,968	37,783
Ohio	11,353	2,133	1,836,554	1,296,450	540,104
Oklahoma	3,451	656	627,032	446,719	180,313
Oregon	3,421	624	545,033	378,474	166,559
Pennsylvania	12,281	2,194	1,816,716	1,262,181	554,535
Rhode Island	1,048	184	156,454	113,520	42,934
South Carolina	4,012	745	666,780	483,725	183,055
South Dakota	755	152	131,037	89,590	41,447
Tennessee	5,689	1,024	916,202	664,393	251,809
Texas	20,852	4,262	3,991,783	2,895,853	1,095,930
Utah	2,233	509	480,255	329,185	151,070
Vermont	609	114	104,559	72,276	32,283
Virginia	7,079	1,276	1,133,994	817,143	316,851
Washington	5,894	1,120	1,003,714	694,750	308,964
West Virginia	1,808	301	291,811	203,475	88,336
Wisconsin	5,364	1,026	877,753	596,439	281,314
Wyoming	494	98	92,105	61,654	30,451
American Samoa	—	—	15,477	11,899	3,578
Guam	—	—	32,951	24,151	8,800
Virgin Islands	—	—	20,866	14,821	6,045

— Data were not available.

[1] Includes a number of prekindergarten students.

SOURCE: U.S. Department of Commerce, Bureau of the Census, *Current Population Reports*, Series P-25, No. 1095 at the national level, SF1-P12 and unpublished data; and U.S. Department of Education, National Center for Education Statistics, Common Core of Data surveys.

Table C.2 **Poverty status of school-age children and children served under Individuals with Disabilities Education Act and Chapter 1, from non-NAEP sources: By state, 1998 and school years 1990–91 through 1999–2000**

	Poverty status of 5- to 17-year-olds: 1998		Children (birth to age 21) served under IDEA[1] and Chapter 1 of the Education Consolidation and Improvement Act, State Operated Programs	
	Number in poverty (in thousands)	Percent in poverty	Number of children: 1999–2000 school year	Percent change: 1990–91 to 1999–2000
Nation	9,167	17.8	6,195,113	30.1
Alabama	156	21.8	99,763	5.1
Alaska	13	9.0	17,495	18.7
Arizona	222	23.6	93,336	63.1
Arkansas	57	13.1	60,864	27.2
California	1,459	22.3	640,815	36.6
Colorado	93	12.5	76,948	34.8
Connecticut	82	13.4	74,722	15.7
Delaware	24	15.7	16,287	13.9
District of Columbia	33	46.0	9,348	48.6
Florida	474	20.5	356,198	50.9
Georgia	377	24.7	164,374	61.2
Hawaii	32	14.5	22,964	74.4
Idaho	50	17.4	29,112	32.2
Illinois	308	12.1	291,221	21.8
Indiana	140	12.6	151,599	32.2
Iowa	73	14.2	71,970	18.6
Kansas	59	13.2	60,036	32.8
Kentucky	118	16.7	91,537	15.3
Louisiana	244	29.8	96,632	31.2
Maine	27	12.0	35,139	25.6
Maryland	66	8.1	111,711	22.4
Massachusetts	163	15.0	165,013	6.7
Michigan	311	14.8	213,404	27.8
Minnesota	130	12.6	107,942	33.4
Mississippi	108	19.3	62,359	2.3
Missouri	136	14.4	134,950	32.4
Montana	42	21.2	19,039	11.1
Nebraska	54	14.8	42,577	30.0
Nevada	49	12.8	35,703	93.6
New Hampshire	34	13.3	28,597	45.5
New Jersey	194	13.2	214,330	18.2
New Mexico	101	23.5	52,346	45.3
New York	848	28.9	434,347	41.3
North Carolina	277	21.3	173,067	40.6
North Dakota	28	17.2	13,612	8.9
Ohio	339	16.0	236,200	15.0
Oklahoma	120	19.9	83,149	26.6
Oregon	121	19.4	73,531	33.3
Pennsylvania	382	18.0	231,175	5.4
Rhode Island	36	20.5	29,895	41.8
South Carolina	129	17.6	103,153	32.6
South Dakota	13	9.2	16,246	8.4
Tennessee	156	14.5	126,732	20.8
Texas	809	20.1	493,850	40.8
Utah	55	11.8	55,389	16.0
Vermont	13	12.2	14,073	14.8
Virginia	92	7.9	161,298	41.5
Washington	118	10.8	116,235	36.1
West Virginia	65	25.7	50,314	16.6
Wisconsin	109	11.5	121,209	39.4
Wyoming	13	13.0	13,307	18.8
American Samoa	—	—	703	93.7
Guam	—	—	2,230	27.4
Virgin Islands	—	—	1,617	21.3

— Data were not available.

[1] Individuals with Disabilities Education Act.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Decennial Census, Minority Economic Profiles, unpublished data; *Current Populaton Reports*, Series P-60, "Poverty in the United States, Money Income of Households, Families, and Persons in the United States, and Income, Poverty, and Valuation of Noncash Benefits, various years, and Money Income in the U.S.: 1999", P60-201; U.S. Department of Education, Office of Special Education and Rehabilitative Services, *Annual Report to Congress on the Implementation of the Individuals with Disabilities Act*, various years.

Table C.3 **Expenditure per pupil, average teacher salary, and pupil/teacher ratio, in public schools, from non-NAEP sources: By state, school years 1998–99 and 2000–01, and fall 1999**

	In public elementary and secondary schools		
	Expenditure per pupil: 1998–99	Estimated average annual salary of teachers: 2000–01	Pupil/teacher ratio: Fall 1999
Nation	$6,508	$42,898	16 [1]
Alabama	5,188	37,956	15 [1]
Alaska	8,404	46,986	17
Arizona	4,672	36,302	19
Arkansas	4,956	34,476	14
California	5,801	48,923	21 [1]
Colorado	5,923	39,284	17
Connecticut	9,318	52,100	14
Delaware	7,706	47,047	15
District of Columbia	9,650	48,651	16 [1]
Florida	5,790	37,824	18
Georgia	6,092	42,216	16
Hawaii	6,081	41,980	17
Idaho	5,066	36,375	18
Illinois	6,762	48,053	16
Indiana	6,772	43,055	17
Iowa	6,243	36,479	15
Kansas	6,015	39,432	14
Kentucky	5,560	37,234	15
Louisiana	5,548	34,253	17
Maine	7,155	36,256	13
Maryland	7,326	44,997	17
Massachusetts	8,260	47,523	13
Michigan	7,432	49,975	18
Minnesota	6,791	40,577	15
Mississippi	4,565	32,957	16
Missouri	5,855	36,764	14
Montana	5,974	32,930	15
Nebraska	6,256	34,175	14
Nevada	5,587	40,172	19
New Hampshire	6,433	38,303	15
New Jersey	10,145	53,281	13
New Mexico	5,440	33,785	16
New York	9,344	50,920	14
North Carolina	5,656	41,167	16
North Dakota	5,442	30,891	14
Ohio	6,627	42,716	16
Oklahoma	5,303	34,434	15
Oregon	6,828	42,333	20
Pennsylvania	7,450	49,500	16
Rhode Island	8,294	48,474	14
South Carolina	5,656	37,327	15
South Dakota	5,259	30,265	14
Tennessee	5,123	37,074	15 [1]
Texas	5,685	38,614	15
Utah	4,210	36,049	22
Vermont	7,541	38,651	12
Virginia	6,350	40,197	14 [1]
Washington	6,110	42,101	20
West Virginia	6,677	35,764	14
Wisconsin	7,527	41,646	14
Wyoming	6,842	34,189	13
American Samoa	2,283	—	19
Guam	—	—	18
Virgin Islands	6,983	—	14

— Data were not available.
[1] Includes imputations for underreporting.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, Revenues and Expenditures for Public Elementary and Secondary Schools, Statistics of State School Systems, and Common Core of Data Surveys; National Education Association, Estimates of School Statistics and unpublished data, 2001.

Acknowledgments

This report is the culmination of the effort of many individuals who contributed their considerable knowledge, experience, and creativity to the NAEP 2002 writing assessment. The 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 Pearson Educational Measurement. Most importantly, NAEP is grateful to the students and school staff who made the assessment possible.

The NAEP 2002 writing assessment was funded through NCES, in the Institute of Education Sciences of the U.S. Department of Education. The Associate Commissioners of Education Statistics, Peggy Carr and Val Plisko, and the NCES staff—Arnold Goldstein, Steven Gorman, Carol Johnson, Andrew Kolstad, Taslima Rahman, and Marilyn Seastrom—worked closely and collegially with the authors to produce this report.

The NAEP project at ETS is directed by Stephen Lazer and John Mazzeo, with assistance from John Barone. Sampling and data collection activities were conducted by Westat under the direction of Rene Slobasky, Nancy Caldwell, Keith Rust, and Dianne Walsh. Printing, distribution, scoring, and processing activities were conducted by Pearson Educational Measurement under the direction of Brad Thayer, Connie Smith, and William Buckles.

Test development activities took place at ETS under the direction of Hilary Persky with assistance from John Morris.

The complex statistical and psychometric activities necessary to report results for the NAEP 2002 reading assessment were directed by Catherine Hombo, Frank Jenkins, and Jiahe Qian, with assistance from Ying Jin.

The extensive data processing and computer programming activities underlying the statistical and psychometric analyses conducted at ETS are under the direction of David Freund, Edward Kulick, Bruce Kaplan, and Steven Isham. Data analyses presented in this report were managed by Scott Davis with assistance from Laura Jerry, Youn-hee Lim, Haiying Liu, Norma Norris, Alfred Rogers, Satwinder Thind, Mike Weiss, and John Willey. The complex database work for this assessment was managed by Katherine Pashley with assistance from Gerry Kokolis.

The reporting process was directed by Jay Campbell with assistance from Wendy Grigg. The design and production of this report were overseen by Loretta Casalaina with assistance from Joseph Kolodey and Rick Hasney. Ming Kuang coordinated the documentation and data checking procedures with assistance from Janice Goodis, Andrea Bergen, and Alice Kass. Arlene Weiner coordinated the editorial procedures with assistance from Trish Hamill. The web version of this report was coordinated by Rick Hasney with assistance from Loretta Casalaina.

Many thanks are due to the numerous reviewers, both internal and external to NCES and ETS. The comments and critical feedback of the following reviewers are reflected in the final version of this report: Arthur Applebee, James Carlson, Ellen Carey, Young Chun, Douglas Cochrane, Mary Crovo, Lawrence Feinberg, Ray Fields, Arnold Goldstein, Steven Gorman, Paula Knepper, Andrew Kolstad, Sarah Murray, Laurence Ogle, Patricia Porter, Taslima Rahman, Linda Shafer, and Marilyn Seastrom.

United States
Department of Education
ED Pubs
8242-B Sandy Court
Jessup, MD 20794-1398

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17