NATIONAL CENTER FOR EDUCATION STATISTICS

TECHNICAL REPORT

NAEP 1996
State Assessment Program in Science

THE NATION'S REPORT CARD naep

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

What is The Nation's Report Card?

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

NAEP is a congressionally mandated project of the National Center for Education Statistics, the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

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

The National Assessment Governing Board

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report of the NAEP 1996 State Assessment Program in Science

Nancy L. Allen
Spencer S. Swinton
Steven P. Isham
Christine A. Zelenak

In collaboration with:

Nada Ballator
Patrick B. Bourgeacq
Charles L. Brungardt
John Burke
Nancy W. Caldwell
James E. Carlson
John R. Donoghue
John J. Ferris
David S. Freund
James L. Green
Penny James
Eugene G. Johnson
Bruce A. Kaplan
John Mazzeo
Norma A. Norris
Ingeborg U. Novatkoski
Patricia E. O'Reilly
Christine Y. O'Sullivan
Katharine E. Pashley
Clyde M. Reese
Alfred M. Rogers
Keith F. Rust
Brent Studer
Lois H. Worthington

January 1998

U.S. Department of Education
Office of Educational Research and Improvement NCES 98-480

Prepared by Educational Testing Service under a cooperative agreement with the National Center for Education Statistics

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

Office of Educational Research and Improvement
Ricky T. Takai
Acting Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

Education Assessment Group
Gary W. Phillips
Associate Commissioner

January 1998

SUGGESTED CITATION
Allen, N.L., Swinton, S.S., Isham, S.P., & Zelenak, C.A. (1997).
Technical report of the NAEP 1996 state assessment program in science.
Washington, DC: National Center for Education Statistics.

FOR MORE INFORMATION
Contact:
Arnold A. Goldstein
202-219-1741

For ordering information on this report, write:

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

or call 1-800-424-1616 (in the Washington, DC, metropolitan area call 202-219-1651) or visit the World Wide Web site: http://nces.ed.gov/naep

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

TECHNICAL REPORT OF THE NAEP 1996 STATE ASSESSMENT PROGRAM IN SCIENCE

TABLE OF CONTENTS

Page

Page

Page

TECHNICAL REPORT OF THE NAEP 1996 STATE ASSESSMENT PROGRAM IN SCIENCE

LIST OF TABLES

Page

Page

TECHNICAL REPORT OF THE NAEP 1996 STATE ASSESSMENT PROGRAM IN SCIENCE

LIST OF FIGURES

Page

Page

ACKNOWLEDGMENTS

The design, development, administration, analysis, and reporting of the 1996 National Assessment of Educational Progress (NAEP) State Assessment program in science was a collaborative effort among staff from the National Center for Education Statistics (NCES), the National Assessment Governing Board (NAGB), the Council of Chief State School Officers (CCSSO), Educational Testing Service (ETS), Westat, Inc., and National Computer Systems (NCS). This report documents the NAEP design, administration, and data analysis procedures used for the State Assessment. It also indicates what technical decisions were made and the rationale behind those decisions. The development of this report and of the State Assessment is the result of the considerable knowledge, experience, creativity, and dedication of many individuals. I would like to acknowledge these individuals for their contribution to NAEP.

The NAEP 1996 science state assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. The Commissioner of Education Statistics, Pascal D. Forgione, Jr., and the NCES staff — Sue Ahmed, Peggy Carr, Arnold Goldstein, Steven Gorman, Larry Ogle, Gary Phillips, Sharif Shakrani, Maureen Treacy, and Alan Vanneman — worked closely and collegially with the authors to produce this report. I would like to especially recognize and thank Shi-Chang Wu and Janis Brown for their significant assistance and support in coordinating the organization and technical review of this report. I would also like to gratefully acknowledge Robert Burton, Andrew Kolstad, Ralph Lee, Marilyn McMillen, and Shi-Chang Wu of NCES, and Jamal Abedi of UCLA's Center for the Study of Evaluation, National Center for Research on Evaluation, for conducting a technical review of this report.

ETS management and the Center for the Assessment of Educational Progress (CAEP) have been very supportive of NAEP's technical work. Special thanks go to Nancy Cole as well as to Henry Braun, Paul Ramsey, Charles Davis, and Juliet Shaffer of ETS research management, and Archie Lapointe, Paul Williams, Steve Lazer, John Mazzeo, Clyde Reese, and Kim Whittington of CAEP. Significant contributions to the project were received from Nada Ballator and Christine O'Sullivan.

The guidance of the NAEP Design and Analysis Committee on technical aspects of NAEP has been outstanding. During the period of analysis of the 1996 data, the members were Sylvia Johnson (chair), Albert Beaton, Jeri Benson, Jeremy Finn, Paul Holland, Huynh Huynh, Gaea Leinhardt, David Lohman, Anthony Nitko, Ingram Olkin, Tej Pandey, Juliet Shaffer, and Hariharan Swaminathan.

The design and data analysis of the 1996 National Assessment was primarily the responsibility of the NAEP research and data analysis staff at ETS with significant contributions from NAEP management, Westat, and NCS staffs. In addition to managing day-to-day data analytic operations, these staff members have made many innovative statistical and psychometric contributions. Major contributions were made by James Carlson, John Donoghue, and Spencer Swinton. Jo-lin Liang served as technical assistant. Eugene Johnson, Robert Mislevy, and Juliet Shaffer provided valuable statistical and psychometric advice.

The Divison of Data Analysis and Technology Research at ETS, under the leadership of John Barone, was responsible for developing the operating systems and carrying out the data analyses. David Freund and Alfred Rogers developed and maintained the large and complex NAEP data management systems, and Katharine Pashley managed database activities. Alfred Rogers developed the production versions of key analysis and scaling systems. Special recognition goes to Phillip Leung and Patricia

O'Reilly for their expertise in the design and development of the NAEP Web site. Thanks also go to David Freund, Steven Isham, Bruce Kaplan, Debbie Kline, and Edward Kulick, for their continuing roles as leaders and developers of innovative solutions to NAEP data analysis challenges. Particular recognition goes to Steven Isham and Lois Worthington, for their substantial contributions to NAEP data analyses, and to Laura Jerry for the computer generated state reports. Many other members of this division made important contributions of their time and talent to NAEP data analyses and analysis software and data products, including John Ferris, Shuyi Hua, Gerry Kokolis, Laura Jenkins, Mike Narcowich, Jennifer Nelson, Norma Norris, Ingeborg Novatkoski, Craig Pizzuti, Steve Szyszkiewicz, Steve Wang, and Xiaohui Wang.

The staff at Westat contributed their talents and efforts in all areas of the sample design and data collection. Particular recognition is due to Rene Slobasky, Nancy Caldwell, Keith Rust, Debby Vivari, and Dianne Walsh for directing the sampling and data collection activities. Thanks are also due to John Burke, James Green, and Penny James.

Critical to the program was the contribution of NCS, which has been responsible for the printing, distribution, and processing of the assessment materials, as well as an increased role in the professional scoring. The leadership roles of Patrick Bourgeacq and Charles Brungardt are especially acknowledged. Thanks also go to Brent Studer.

Carol Crowley, Carol Manikin, Sharon Stewart, and Terri Stirling are acknowledged for their administrative assistance during the preparation of this report and in other related activities.

Carol Errickson also deserves special thanks for coordinating the design of the report's cover.

Special recognition and appreciation go to Christine Zelenak, editor of this report, for organizing, scheduling, editing, motivating, and ensuring the cohesiveness and correctness of the final report.

Finally, NAEP is grateful to the students and school staff members who participated in the assessment. Without their efforts, there would be no assessment.

Nancy L. Allen
Director of Data Analysis and Scaling
NAEP Research, ETS
November 1997

Chapter 1

OVERVIEW: THE DESIGN, IMPLEMENTATION, AND ANALYSIS OF THE 1996 STATE ASSESSMENT PROGRAM IN SCIENCE[1]

Nancy L. Allen and John Mazzeo
Educational Testing Service

1.1 OVERVIEW

In April 1988, Congress reauthorized the National Assessment of Educational Progress (NAEP) and added a new dimension to the program — voluntary state-by-state assessments on a trial basis in 1990 and 1992, in addition to continuing the national assessments that NAEP had conducted since its inception. In 1994, Congress authorized a third Trial State Assessment for administration in 1994. It should be noted that the word *trial* in Trial State Assessment refers to the Congressionally mandated trial to determine whether such assessments can yield valid, reliable state representative data. Enough experience had been gained for Congress to authorize State Assessments, rather than Trial State Assessments, to be conducted in 1996. In this report, we will refer to the voluntary state-by-state assessment program as the State Assessment program. The State Assessment program, which is designed to provide representative data on achievement for participating jurisdictions, is distinct from the assessment designed to provide nationally representative data, referred to in this report as the national assessment. (This terminology is also used in all other reports of the 1996 assessment results.) All instruments and procedures used in the 1990, 1992, 1994, and 1996 state and national assessments were previously piloted in field tests conducted in the year prior to each assessment.

The 1990 Trial State Assessment program collected information on the mathematics knowledge, skills, understanding, and perceptions of a representative sample of eighth-grade students in public schools in 37 states, the District of Columbia, and two territories. The second phase of the Trial State Assessment program, conducted in 1992, collected information on the mathematics knowledge, skills, understanding, and perceptions of a representative sample of fourth- and eighth-grade students and the reading skills and understanding of a representative sample of fourth-grade students in public schools in 41 states, the District of Columbia, and two territories.

The 1994 Trial State Assessment program once again assessed the reading skills and understanding of representative samples of fourth-grade students, this time in 44 participating jurisdictions. The 1994 program broke new ground in two ways. The 1994 NAEP authorization called for the assessment of samples of both public- and nonpublic-school students. Thus, for the first time in NAEP, jurisdiction-level samples of students from Catholic schools, other religious schools and private schools, Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS), and Bureau of Indian Affairs (BIA) schools were added to the Trial State Assessment program. Second, samples of students from the Department of Defense Dependents Schools Office of Dependents Education (DoDDS) schools participated as a

[1]Nancy L. Allen is the Director of Data Analysis and Scaling, NAEP Research, Educational Testing Service. John Mazzeo is the Director of NAEP Reporting, Educational Testing Service.

jurisdiction, along with the states and territories that have traditionally had the opportunity to participate in the Trial State Assessment program.

The 1996 State Assessment program, described in this report, collected information on the science knowledge, skills, understanding, and perceptions of a representative sample of eighth-grade students in the jurisdictions shown in Table 1-1; Department of Defense Education Activity (DoDEA) school students were assessed at both grades 4 and 8. The grade 4 assessment of DoDEA students was a special assessment supported by NCES. In addition, grade 4 and grade 8 students were assessed for a third time in mathematics (see the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*, Allen, Jenkins, Kulick, & Zelenak, 1996).

A special feature of the 1996 State Assessments was the introduction of new rules for student inclusion in NAEP assessments. Half of the schools selected for participation in the 1996 assessment used the old inclusion rules to determine whether students should be included in the assessment and the other half used the new inclusion rules. In addition to the two groups of schools using the old and new inclusion rules without offering students special testing accommodations, the 1996 national assessment included a third group of schools that used the new inclusion rules and offered students within those schools accommodations to the standard NAEP administration procedures.

The accommodations provided by NAEP in the national assessments were meant to match those specified in the student's individualized education plan (IEP) or those ordinarily provided in the classroom for testing situations. The most common accommodation was extended time. In the State Assessment, no special accommodations were offered.

The new inclusion rules are applied only when a student has been categorized in his or her IEP as a student with disabilities (SD) or as a student with limited English proficiency (LEP); all other students are asked to participate in the assessment. For this reason, the sample of students that were selected for most analysis and reporting purposes for science consisted of students who were not categorized as SD or LEP students and students from the schools using new inclusion rules that were categorized as SD or LEP. The students who were not categorized as SD or LEP were from all schools no matter which set of inclusion rules were used. The advantage of this reporting sample is that it makes use of most of the data from the assessment and begins a science trend line for the State Assessment program with the new inclusion rules.

Special analyses that used the national science and mathematics assessment data to compare the old and new inclusion rules and examine the effect of offering testing accommodations, indicated little difference in proportions of students included in the assessment using the old and new inclusions. More students were included in the assessment when they were offered accommodations; however, a portion of students who would have participated in the assessment under standard conditions were assessed with accommodations when they were offered. A result of this is that fewer students were assessed under standard conditions when accommodations were offered.

Table 1-1 lists the jurisdictions that participated in the 1996 State Assessment program. Over 125,000 students participated in the 1996 State Assessment in science in the jurisdictions shown. Students were administered the same assessment booklets that were used in NAEP's 1996 national science assessment.

The 1996 NAEP science framework and assessment specifications were developed for NAEP through a consensus project conducted by the Council of Chief State School Officers (CCSSO) under funding from the National Assessment Governing Board (NAGB). During this development process, input and reactions were continually sought from a wide range of educators and professionals both within the field of science and external to it. Hence, for grade 8, the assessment provides the first opportunity to report jurisdiction-level data for a NAEP science instrument for those states and territories that participated in the 1996 State Assessment program. In addition, questionnaires completed by the students, their science teachers, and principals or other school administrators provided an abundance of contextual data within which to interpret the science results.

Table 1-1

Jurisdictions Participating in the 1996 State Assessment Program in Science[1]

Jurisdictions			
Alabama	Georgia	Mississippi	Rhode Island
Alaska	Guam	Missouri	South Carolina
Arizona	Hawaii	Montana	Tennessee
Arkansas	Indiana	Nebraska	Texas
California	Iowa	Nevada	Utah
Colorado	Kentucky	New Hampshire	Vermont
Connecticut	Louisiana	New Jersey	Virginia
Delaware	Maine	New Mexico	Washington
DoDEA/DDESS	Maryland	New York	West Virginia
DoDEA/DoDDS	Massachusetts	North Carolina	Wisconsin
District of Columbia	Michigan	North Dakota	Wyoming
Florida	Minnesota	Oregon	

[1]The 1996 State Assessment in science was conducted at grade 8 only, although Department of Defense Education Activity (DoDEA) school students were also assessed as part of a separate special assessment.

The purpose of this report is to provide technical information about the 1996 State Assessment in science. It provides a description of the design for the State Assessment and gives an overview of the steps involved in the implementation of the program from the planning stages through to the analysis and reporting of the data. As stated previously, the 1996 State Assessment in science was conducted at grade 8 only, although, as part of a special assessment, DoDEA students in grade 4 were also assessed. The report describes in detail the development of the cognitive and background questions, the field procedures, the creation of the database and data products for analysis, and the methods and procedures used for sampling, analysis, and reporting. It does not provide the results of the assessment — rather, it provides information on how those results were derived.

This report is one of several documents that provide technical information about the 1996 State Assessment. For those interested in performing their own analyses of the data, this report and the user guide for the secondary-use data should be used as primary sources of information about NAEP (O'Reilly, Zelenak, Rogers, & Kline, 1997). Information for lay audiences is provided in the procedural appendices to the science subject-area reports; theoretical information about the models and procedures used in NAEP can be found in the special NAEP-related issue of the *Journal of Educational Statistics* (Summer 1992/Volume 17, Number 2) and in previous national technical reports. Further, the *Science Framework for the 1996 National*

Assessment in Educational Progress includes a discussion of the processes and specifications by which the framework was developed (National Assessment Governing Board, 1993). For more information about the science assessment and the characteristics of the items in the assessment, see *The NAEP 1996 Technical Report* (Allen, Carlson, & Zelenak, 1998).

Under a cooperative agreement with the National Center for Education Statistics (NCES), Educational Testing Service (ETS) was responsible for the development, analysis, and reporting of the 1996 NAEP programs, including the State Assessment. ETS was responsible for overall management of aspects of the programs as well as for development of the overall design, the items and questionnaires, data analysis, and reporting. National Computer Systems (NCS) was a subcontractor to ETS on both the national and State NAEP programs. NCS was responsible for printing, distribution, and receipt of all assessment materials, and for data processing, scanning, and professional scoring. All aspects of sampling and field operations for both the national and State Assessments were the responsibility of Westat, Inc. The National Center for Education Statistics (NCES) contracted directly with Westat for these services for the national and state assessments.

This technical report provides information about the technical bases for a series of reports that have been prepared for the 1996 State Assessment program in science. They include:

- A *State Report* for each participating jurisdiction that describes the science scale scores of the eighth-grade public- and nonpublic-school students in that jurisdiction and relates their scale scores to contextual information about science policies and instruction.

- The *NAEP 1996 Science Report Card for the Nation and the States*, which provides both public- and nonpublic-school data for major NAEP reporting subgroups for all of the jurisdictions that participated in the State Assessment program, as well as selected results from the 1996 national science assessment.

- The *Cross-State Data Compendium for the NAEP 1996 Science Assessment*, which includes jurisdiction-level results for all the demographic, instructional, and experiential background variables included in the *Science Report Card* and *State Report*.

- A *Data Almanac* for each jurisdiction, distributed only in electronic form, that contains a detailed breakdown of the science scale-score data according to the responses to the student, teacher, and school questionnaires for the public school, nonpublic school, and combined populations as a whole and for important subgroups of the public-school population. There are six sections to each almanac:

  ⇒ *The Distribution Data Section* provides selected percentiles for the public school, nonpublic school, and combined populations and for the standard demographic subgroups of the public-school population for the composite scale and each field of science scale.[2]

[2] Scales were created for the three fields of science: earth science, physical science, and life science.

⇒ *The Student Questionnaire Section* provides a breakdown of the composite scale score data according to the students' responses to questions in the three student questionnaires included in the assessment booklets.

⇒ *The Teacher Questionnaire Section* provides a breakdown of the composite scale score data according to the teachers' responses to questions in the science teacher questionnaire.

⇒ *The School Questionnaire Section* provides a breakdown of the composite scale score data according to the principals' (or other administrators') responses to questions in the school characteristics and policies questionnaire.

⇒ *The Scale Section* provides a breakdown of selected items from the questionnaires according to each of the scales measuring the fields of science in the assessment.

⇒ *The Science Item Section* provides the response data for each science item in the assessment.

The state reports and the *Science Report Card* will be available on the World Wide Web as they are publicly released; the almanacs will be placed on the web about a month after they are released on CD-ROM.

Organization of the Technical Report

This chapter provides a description of the design for the State Assessment in science and gives an overview of the steps involved in implementing the program from the planning stages to the analysis and reporting of the data. The chapter summarizes the major components of the program, with references to later chapters for more details. Because of the close relationship between the grade 8 State Assessment in science and a science assessment of grade 4 students in DoDEA schools, this document also provides technical information about the special assessment of grade 4 DDESS and DoDDS students in science. This special assessment was directly contracted for by NCES. The organization of this chapter, and of the report, is as follows:

- Section 1.2 provides an overview of the design of the 1996 State Assessment program in science.

- Section 1.3 summarizes the development of the science objectives and the development and review of the items written to measure those objectives. Details are provided in Chapter 2.

- Section 1.4 discusses the assignment of the cognitive items to assessment booklets. An initial discussion is provided of the complex spiral design that was used to assign cognitive items to assessment booklets and assessment booklets to individuals. A more complete description is provided in Chapter 2.

- Section 1.5 outlines the sampling design used for the 1996 State Assessment program in science. A fuller description is provided in Chapter 3.

- Section 1.6 summarizes the field administration procedures, including securing school cooperation, training administrators, administering the assessment, and conducting quality control. Further details appear in Chapter 4.

- Section 1.7 describes the flow of the data from their receipt at NCS through data entry, professional scoring, and entry into the ETS/NAEP database for analysis, and the creation of data products for secondary users. Chapters 5 and 6 provide a detailed description of the process.

- Section 1.8 provides an overview of the data obtained from the 1996 State Assessment program in science.

- Section 1.9 summarizes the procedures used to weight the assessment data and to obtain estimates of the sampling variability of subpopulation estimates. Chapter 7 provides a full description of the weighting and variance estimation procedures.

- Section 1.10 describes the initial analyses performed to verify the quality of the data in preparation for more refined analyses, with details given in Chapter 9.

- Section 1.11 describes the item response theory scales and the overall science composite that were created for the primary analysis of the State Assessment data. Further discussion of the theory and philosophy of the scaling technology appears in Chapter 8, with details of the scaling process in Chapter 9.

- Section 1.12 provides an overview of the linking of the scaled results from the State Assessment to those from the national science assessment. Details of the linking process appear in Chapter 9.

- Section 1.13 describes the reporting of the assessment results, with further details supplied in Chapter 10.

- Section 1.14 indicates some of the features of the special assessment of grade 4 DoDEA students in science. Further information is presented in each chapter of this report, as appropriate.

- Appendices A through F include a list of the participants in the objectives and item development process, a summary of the participation rates, a list of the conditioning variables, the IRT parameters for the science items, the reporting subgroups, and composite and derived common background and reporting variables.

1.2 DESIGN OF THE STATE ASSESSMENT IN SCIENCE

The major aspects of the design for the State Assessment in science included the following:

- Participation at the jurisdiction level was voluntary.

- Eighth-grade students from public and nonpublic schools were assessed. Nonpublic schools included Catholic schools, other religious schools, private schools, DoDEA schools,[3] and BIA schools. Separate representative samples of public and nonpublic schools were selected in each participating jurisdiction and students were randomly sampled within schools. The sizes of a jurisdiction's nonpublic-school samples were proportional to the percentage of grade-level students in that jurisdiction attending such schools.

- The eighth-grade science student booklets used for the 1996 NAEP State Assessment, and included as part of the 1996 national NAEP instrument contained multiple-choice, short-constructed response, and extended-constructed response cognitive items. The total pool of science items was divided into 15 blocks of items, each 30 minutes long, at each grade level.

- A complex form of matrix sampling using *spiraling* of assessment booklets was used. With spiraling, students in an assessment session received different booklets, which provides for greater science content coverage than would have been possible had every student been administered the identical set of items, without imposing an undue testing burden on the student.

- Background questionnaires given to the students, the students' science teachers, and the principals or other administrators provided a variety of contextual information. The background questionnaires for the State Assessment program were identical to those used in the national assessments.

- The assessment time for each student was approximately 103 minutes. Each assessed student was assigned a science booklet that contained two 5-minute background questionnaires, one 3-minute motivation questionnaire, and three of the 15 blocks containing science items requiring 30 minutes each. Thirty-seven different booklets were assembled.

---

[3]Students from two of the DDESS schools were included as part of the State Assessment and in the special assessment of DoDEA schools. In these cases, the DDESS school ID was replaced with the state ID.

- The assessments took place in the five-week period between January 29 and March 4, 1996. One-fourth of the schools in each jurisdiction were to be assessed each week throughout the first four weeks; however, due to severe weather throughout much of the country, the fifth week was used for regular testing as well as for makeup sessions.

- Data collection was, by law, the responsibility of each participating jurisdiction. Security and uniform assessment administration were high priorities. Extensive training of State Assessment personnel was conducted to assure that the assessment would be administered under standard, uniform procedures. For jurisdictions that had participated in previous NAEP state assessments, 25 percent of both public- and nonpublic-school assessment sessions were monitored by Westat staff. For the jurisdictions new to NAEP, 50 percent of both public- and nonpublic-school sessions were monitored.

1.3 DEVELOPMENT OF SCIENCE OBJECTIVES, ITEMS, AND BACKGROUND QUESTIONS

The science framework for the 1996 NAEP was produced under the auspices of the NAGB. The consensus process was managed by the Council of Chief State School Officers (CCSSO) who worked with the National Center for Improving Science Education and the American Institutes for Research. Items were developed that were aligned with the specifications described in the framework and were extensively reviewed by specialists in science, measurement, and bias/sensitivity, as well as by government officials and state representatives.

A Planning Committee was established to identify goals and objectives and to produce the framework. This Planning Committee met monthly from November 1990 through April 1991 and was joined in the first meeting and final meeting by the Steering Committee, which reviewed and reacted to all framework drafts. During this development process, input and reactions were continually sought from a wide range of members both within the field of science and external to it.

The framework for the 1996 science assessment is represented as a matrix with two dimensions represented by three fields of science (*earth science*, *physical science*, and *life science*) and three elements of knowing and doing science (*conceptual understanding*, *scientific investigation*, and *practical reasoning*). In addition, there are two subcategories that describe science, nature of science, and themes (National Assessment Governing Board, 1993).

Chapter 2 includes specific details about developing the objectives and items for the State Assessment. Further information about the items in the assessment are available in research papers presented at recent annual meetings of the American Educational Research Association and the National Council on Measurement in Education (e.g., Carlson, 1996; Yepes-Baraya & Allen, 1996; Worthington & Donoghue, 1997, Yepes-Baraya, 1997; Allen & Liang, 1997; and Tatsuoka, 1997) and in an upcoming science focus report.

1.4 ASSESSMENT INSTRUMENTS

The assembly of cognitive items into booklets and their subsequent assignment to assessed students was determined by a complex design with spiraled administration. Details of this design are provided in Chapter 2. Every student was asked to complete a hands-on performance task as well as two other cognitive blocks of items in a paper and pencil format. Some students received paper and pencil blocks based on a science theme. In addition to the student assessment booklets, three other instruments provided data relating to the assessment — a science teacher questionnaire, a school characteristics and policies questionnaire, and an SD/LEP student questionnaire.

The *student assessment booklets* contained five sections and included both cognitive and noncognitive questions. In addition to three 30-minute sections of cognitive questions, each booklet included two 5-minute sets of general and science background items designed to gather contextual information about students, their experiences in science, and their attitudes toward the subject, and one 3-minute section of motivation questions designed to gather information about the student's level of motivation while taking the assessment.

The *teacher questionnaire* was administered to the eighth-grade science teachers of the students participating in the assessment. The questionnaire consisted of three sections and took approximately 20 minutes to complete. The first section focused on the teacher's general background and experience; the second, on the teacher's background related to science; and the third, on classroom information about science instruction.

The *school characteristics and policies questionnaire* was given to the principal or other administrator in each participating school and took about 20 minutes to complete. The questions asked about the principal's background and experience; school policies, programs, and facilities; and the demographic composition and background of the students and teachers.

The *SD/LEP student questionnaire* was completed by the teachers of those students who were selected to participate in the State Assessment sample but who were classified as students with disabilities (SD) or were categorized as having limited English proficiency (LEP). Some of these students did not participate in the assessment because they were determined by the school personnel to be unable to participate, using inclusion rules provided by NAEP; others did participate in the assessment because they were determined to be able to participate by meeting the specifications in the inclusion rules. Each questionnaire took approximately three minutes to complete and asked about the student and the special programs in which the student participated.

Further information on the assessment instruments can be found in Chapter 2.

1.5 THE SAMPLING DESIGN

The target populations for the State Assessment program in science consisted of eighth-grade students enrolled in public and nonpublic schools. The public- and nonpublic-school samples in each jurisdiction were designed to produce aggregate estimates for the jurisdiction and for selected subpopulations (depending upon the size and distribution of the various subpopulations within the jurisdiction), and also to enable comparisons to be made, at the jurisdiction level, between administration of assessment tasks with and without monitoring.

The representative sample of public-school eighth-grade students assessed in the State Assessment came from about 100 schools in most jurisdictions. However, if a jurisdiction had fewer than 100 schools with an eighth grade, all or almost all schools were asked to participate. If a jurisdiction had smaller numbers of students in each school than expected, more than 100 schools were selected for participation. The public schools were stratified by urbanization, percentage of Black and Hispanic students enrolled, and median household income within the ZIP code area of the school.

The nonpublic-school samples differed in size across the jurisdictions, with the number of schools selected proportional to the nonpublic-school enrollment within each jurisdiction. Typically, about 20 to 25 nonpublic schools (per grade) were included for each jurisdiction. The nonpublic schools were stratified by type of control (Catholic, private/other religious, other nonpublic), metro status, and enrollment size per grade.

In most jurisdictions, up to 30 students were selected from each school, with the aim of providing an initial target sample size of approximately 3,000 public-school students per jurisdiction. The student sample size of 30 for each school was chosen to ensure that at least 2,000 public-school students participated from each jurisdiction allowing for school nonresponse, exclusion of students, inaccuracies in the measures of enrollment, and student absenteeism from the assessment. In jurisdictions with fewer schools, larger numbers of students per school were often required to ensure target samples of roughly 3,000 students. In certain jurisdictions, all eligible eighth-grade students were targeted for assessment. The overall student sample size for nonpublic schools was much smaller than the approximate 2,000 students from public schools that were assessed.

Students within a school were sampled from lists of eighth-grade students. The decisions to exclude students from the assessment were made by school personnel, in one of two ways. The students in one group of schools were excluded using the inclusion rules used in previous assessments and students in a second group of schools were excluded on the basis of inclusion rules that were new for the 1996 assessment. The new inclusion rules are meant to be clearer, more easily followed, and closer to inclusion rules used in testing programs administered by school districts or state departments of education. In the 1996 national assessments, students in a third group of schools were excluded using the new inclusion rules, but SD and LEP students in these schools were offered special accommodations to the standard NAEP administration procedures. In the State Assessment, no special accommodations were offered. Each excluded student in the State Assessment was carefully accounted for to estimate the percentage of the state population deemed unassessable and the reasons for exclusion, no matter which school the student attended.

Chapter 3 describes the various aspects of selecting the sample for the 1996 State Assessment — selection of schools for use of the differing inclusion criteria, the construction of the public- and nonpublic-school frames, the stratification processes, the updating of the school frames with new schools, the actual sample selection, and the sample selection for the field test.

1.6 FIELD ADMINISTRATION

The administration of the 1996 program and the 1995 field test required collaboration between staff in the participating jurisdictions and schools and the NAEP contractors, especially Westat, the field administration contractor. The purpose of the field test conducted in 1995 was to try out new science items including those associated with science themes and the hands-on science performance tasks.

Each jurisdiction volunteering to participate in the 1995 field test or in the 1996 State Assessment program was asked to appoint a state coordinator as liaison between NAEP staff and the participating schools. In addition, Westat hired and trained a supervisor for each jurisdiction and six field managers, each of whom was assigned to work with groups of jurisdictions. The state supervisors were responsible for working with the state coordinators, overseeing assessment activities, training school district personnel to administer the assessment, and coordinating the quality-control monitoring efforts. Each field manager was responsible for working with the state coordinators of seven to eight jurisdictions and supervising the state supervisors assigned to those jurisdictions. An assessment administrator was responsible for preparing for and conducting the assessment session in one or more schools. These individuals were usually school or district staff and were trained by Westat. Westat also hired and trained three to five quality control monitors in each jurisdiction. For jurisdictions that had previously participated in the State Assessment program, 25 percent of the public- and nonpublic-school sessions were monitored. For jurisdictions new to the program, 50 percent of all sessions were monitored. During the field test, the state supervisors monitored all sessions.

Chapter 4 describes the procedures for obtaining jurisdiction cooperation and provides details about the field activities for both the field test and 1996 State Assessment program. Chapter 4 also describes the planning and preparations for the actual administration of the assessment, the training and monitoring of the assessment sessions, and the responsibilities of the state coordinators, state supervisors, assessment administrators, and quality control monitors.

1.7 MATERIALS PROCESSING, PROFESSIONAL SCORING, AND DATABASE CREATION

Upon completion of each assessment session, school personnel shipped the assessment booklets and forms to NAEP contractor NCS for professional scoring, entry into computer files, and checking. The files were then sent to ETS for creation of the database. Chapter 5 describes the printing, distribution, receipt, processing, and final disposition of the 1996 State Assessment materials.

The volume of collected data and the complexity of the State Assessment processing design, with its spiraled distribution of booklets, as well as the concurrent administration of this assessment and the national assessments, required the development and implementation of flexible, innovative processing programs, and a sophisticated Process Control System. This system, described in Chapter 5, allowed an integration of data entry and workflow management systems that included carefully planned and delineated editing, quality control, and auditing procedures.

Chapter 5 also provides information about scoring procedures and rater reliability. Further and more detailed information about these topics are provided in *The NAEP 1996 Technical Report* (Allen, Carlson, & Zelenak, 1998). The data transcription and editing procedures used to generate the electronic files containing various assessment information, including the sampling weights required to make valid statistical inferences about the population from which the State Assessment sample was drawn, are also described in Chapter 5. Before any analysis could begin, the data from these files underwent a quality control check at ETS. The files were then merged into a comprehensive, integrated database. Chapter 6 describes the transcribed data files, the procedure of merging them to create the State Assessment database, the results of the quality control process, and the procedures used to create data products for use in secondary research.

1.8 THE 1996 STATE ASSESSMENT DATA

The basic information collected from the State Assessment in science consisted of the responses of the assessed students to the 195 science exercises at grade 8. To limit the assessment time for each student to about 103 minutes, a complex variant of matrix and spiraled administration was used to assign a subset of the full exercise pool to each student. The set of items was divided into 15 unique blocks, each requiring 30 minutes for completion. Each assessed student received a booklet containing three of the 15 blocks according to a complex design that ensured that each block was administered to a representative sample of students within each jurisdiction. The data also included responses to the background questionnaires (described in Section 1.4 of this chapter and in Chapter 2).

The national data to which the State Assessment results were compared came from nationally representative samples of public- and nonpublic-school students in the eighth grade. These samples were part of the full 1996 national science assessment in which nationally representative samples of students in public and nonpublic schools were assessed from three grade cohorts: fourth-, eighth-, and twelfth-grade students.

The assessment instruments used in the State Assessment were also used in the eighth-grade national assessment and were administered using almost identical procedures in both assessments. The time of testing for the state assessments (January 29-March 4, 1996) occurred within the time of testing of the national assessment (January 3-April 5, 1996). However, the state assessment differed from the national assessment in one important regard: Westat staff collected the data for the national assessment while, in accordance with the NAEP legislation, data collection activities for the State Assessment were the responsibility of each participating jurisdiction. These activities included ensuring the participation of selected schools and students, assessing students according to standardized procedures, and observing procedures for test security. To provide quality control of the State Assessment, a random half of the administrations in jurisdictions participating in a State Assessment for the first time was monitored; 25 percent of the administrations in other jurisdictions were monitored.

1.9 WEIGHTING AND VARIANCE ESTIMATION

A complex sample design was used to select the students to be assessed in each of the participating jurisdictions. The properties of a sample from a complex design are very different from those of a simple random sample in which every student in the target population has an equal chance of selection and every combination of students of the size of the sample has an equal chance of selection. The properties of the sample from the complex State Assessment design were taken into account in 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. These weights also included adjustments for school and student nonresponse. All population and subpopulation characteristics based on the State Assessment data used sampling weights in their estimation. Chapter 7 provides details on the computation of these weights.

In addition to deriving appropriate estimates of population characteristics, it is essential to obtain appropriate measures of the degree of uncertainty of those statistics. One component of uncertainty is a result of sampling variability, which measures the dependence of the results on the particular sample of students actually assessed. Because of the effects of cluster selection (schools are selected first, then students are selected within those schools), observations made on different students cannot be assumed to be independent of each other (and, in fact, are generally positively correlated). As a result, classical variance estimation formulas will produce incorrect results. Instead, a variance estimation procedure that takes the characteristics of the sample into account was used for all analyses. This procedure, called *jackknife variance estimation*, is discussed in Chapter 7 and described more fully in *The NAEP 1994 Technical Report* (Allen, Kline, & Zelenak, 1996).

Jackknife variance estimation provides a reasonable measure of uncertainty for any statistic based on values observed without error. Statistics such as the average proportion of students correctly answering a given question meet this requirement, but other statistics based on estimates of student science performance, such as the average science scale score of a subpopulation, do not. Because each student typically responds to relatively few items within a particular field of science, there exists a nontrivial amount of imprecision in the measurement of the proficiency of a given student. This imprecision adds an additional component of variability to statistics based on estimates of individual scale scores. The estimation of this component of variability is discussed in Chapter 8.

1.10 PRELIMINARY DATA ANALYSIS

After the computer files of student responses were received from NCS, all cognitive and noncognitive items were subjected to an extensive item analysis. Each block of cognitive items was subjected to item analysis routines, which yielded for each item the number of respondents, the percentage of responses in each response category for an item, the percentage who omitted the item, the percentage who did not reach the item, and the correlation between the item score and the item block score. In addition, the item analysis program provided summary statistics for each block of items, including a reliability (internal consistency) coefficient. These analyses were used to check on the scoring of the items, to verify the appropriateness of the difficulty level of the items, and to check for speededness. The results also were reviewed by knowledgeable project staff in search of aberrations that might signal unusual results or errors in the database.

Tables of the weighted percentages of students with responses in each category of each cognitive and background item were created and distributed to each jurisdiction. Additional analyses comparing the data from the monitored sessions with those from the unmonitored sessions were conducted to determine the comparability of the assessment data from the two types of administrations. Differential item functioning (DIF) analyses using national assessment data were carried out to identify items new to the assessment that were differentially difficult for various subgroups and to reexamine such items with respect to their fairness and their appropriateness for inclusion in the scaling process. Further details of the preliminary analyses conducted on the data appear in Chapter 9.

1.11 SCALING THE ASSESSMENT ITEMS

The primary analysis and reporting of the results from the State Assessment program used item response theory (IRT) scale-score models. Scaling models quantify a respondent's tendency to provide correct answers to the domain of items contributing to a scale as a function of a parameter called proficiency, estimated by a scale score. The scale scores can be viewed as a summary measure of performance across the domain of items that make up the scale. Three distinct IRT models were used for scaling: 1) three-parameter logistic models for multiple-choice items; 2) two-parameter logistic models for short constructed-response items that were scored correct or incorrect; and 3) generalized partial-credit models for short and extended constructed-response items that were scored on a multipoint scale. Chapter 8 provides an overview of the scaling models used. Further details on the application of these models are provided in Chapter 9.

A series of scales were created for the State Assessment to summarize students' science performance. These scales were defined identically to those used for the scaling of the national NAEP eighth-grade science data. Three fields of science scales, based on the paradigm described in Chapter 2, were created to correspond to the following areas: *earth science*, *physical science*, and *life science*. Although the items comprising each scale were identical to those used for the national program, the item parameters for the State Assessment scales were estimated from the combined data from all jurisdictions participating in the State Assessment. Item parameter estimation was based on an item calibration sample consisting of an approximately 25 percent sample of all the available data. To ensure equal representation in the scaling process, each jurisdiction was equally represented in the item calibration sample. Chapter 9 provides further details about the item parameter estimation.

The fit of the IRT model to the observed data was examined within each scale by comparing the estimates of the empirical item characteristic functions with the theoretic curves. For multiple-choice and dichotomously-scored constructed response items, nonmodel-based estimates of the expected proportions of correct responses to each item for students with various levels of scale scores were compared with the fitted item response curve; for partial-credit polytomously-scored constructed-response items, the comparisons were based on the expected proportions of students with various levels of scale scores who achieved each item score level. In general, the item-level results were well fit by the scaling models.

Using the item parameter estimates, estimates of various population statistics were obtained for each jurisdiction. The NAEP methods use random draws ("plausible values") from scale score distributions for each student to compute population statistics. Plausible values are not optimal individual student scale scores; instead, they serve as intermediate values to be used in estimating population characteristics. Under the assumptions of the scaling models, these population estimates will be consistent, in the sense that the estimates approach the model-based population values as the sample size increases. This would not be the case for population estimates obtained by aggregating optimal individual scale scores. Chapter 8 provides further details on the computation and use of plausible values.

In addition to the plausible values for each scale, a composite score scale of the three fields of science scales was created as a measure of overall science proficiency. This composite was a weighted average of the three fields of science scales in which the weights were proportional to the relative importance assigned to each field as specified in the science objectives. The definitions of the composites for the State Assessment program at grade 8 were identical to those used for the national eighth-grade science assessments.

1.12 LINKING THE STATE RESULTS TO THE NATIONAL RESULTS

A major purpose of the State Assessment program was to allow each participating jurisdiction to compare its 1996 results with the nation as a whole and with the region of the country in which that jurisdiction is located. For meaningful comparisons to be made between each of the State Assessment jurisdictions and the relevant national sample, results from these two assessments had to be expressed in terms of a similar system of scale units.

The results from the State Assessment program were linked to those from the national assessment through linking functions determined by comparing the results for the aggregate of all eighth-grade students assessed in the State Assessment with the results for students of the matching grade within a subsample (the National Linking sample) of the national NAEP sample. The National Linking sample for a given grade is a representative sample of the population of all grade-eligible public-school students within the aggregate of the 44 participating states and the District of Columbia (excluding Guam and the two DoDEA jurisdictions). Specifically, the grade 8 National Linking sample consists of all eighth-grade students in public schools in the states and the District of Columbia who were assessed in the national science assessment.

A linear equating within each scale was used to link the results of the State Assessment to the national assessment. For each scale, the adequacy of linear equating was evaluated by comparing the distribution of science scale scores based on the aggregation of all assessed students from the participating states and the District of Columbia with the equivalent distribution based on the students in the National Linking sample. In the estimation of these distributions, the students were weighted to represent the target population of public-school students in the aggregation of the states and the District of Columbia. If a linear equating was adequate, the distribution for the aggregate of states and the District of Columbia and that for the National Linking sample would have, to a close approximation, the same shape, in terms of the skewness, kurtosis, and higher moments of the distributions. The only differences in the distributions allowed by linear equating are in the means and variances. This has been found to be the case for the 1996 State Assessment program.

Each field of science scale was linked by matching the mean and standard deviation of the scale score averages across all eighth-grade students in the State Assessment to the corresponding scale mean and standard deviation across all students in the eighth-grade National Linking sample. Further details of the linking are given in Chapter 9.

1.13 REPORTING THE STATE ASSESSMENT RESULTS

Each jurisdiction in the State Assessment received a summary report providing its results with accompanying text and tables and national and regional comparisons. These reports were generated by a computerized report-generation system for which graphic designers, statisticians, data analysts, and report writers collaborated to develop shells of the reports in advance of the analysis. These prototype reports were provided to State Education Agency personnel for their reviews and comments. The results of the data analysis were then automatically incorporated into the reports, which display tables and graphs of the results and interpretations of those results, including indications of subpopulation comparisons of statistical and substantive significance.

Each report contains state-level estimates of average scale score, both for the state as a whole and for categories of the key reporting variables: gender, race/ethnicity, level of parental education, and type of location. Results are presented for each science scale score and for the overall science composite scale score. Results are also reported for a variety of other subpopulations based on variables derived from the student, teacher, and school questionnaires. Standard errors are included for all statistics.

A second report, the *NAEP 1996 Science Report Card for the Nation and the States*, highlights key assessment results for the nation and summarizes results across the jurisdictions participating in the assessment. This report contains composite scale-score results (scale-score means, etc.) for the nation, each of the four regions of the country, and each jurisdiction participating in the State Assessment, both overall and by the primary reporting variables. In addition, overall results are reported for each of the fields of science scales.

The third type of summary report is entitled *Cross-State Data Compendium for the NAEP 1996 Science Assessment*. Like the *Report Card*, the *Compendium* reports results for the nation and for all of the jurisdictions participating in the State Assessment. The *Compendium* contains most of the tables included in the *Report Card* plus additional tables that provide composite scale-score results for a large number of secondary reporting variables.

The fourth type of summary report is a six-section almanac. One section of the almanac includes information about the percentages of students at or above the three composite scale achievement levels (and below basic). Three of the sections of the almanac present analyses based on responses to each of the questionnaires (student, science teacher, and school) administered as part of the State Assessment. Another section of the almanac, the scale section, reports scale score means and associated standard errors for the three fields of science scales. Results in this section are also reported for the total group in each jurisdiction, as well as for select subgroups of interest. The final section of the almanac, the "p-value" section, provides the total-group proportion of correct responses to each cognitive item included in the assessment.

The production of the state reports, *Science Report Card*, *Data Compendium*, and the almanacs required a large number of decisions about a variety of data analysis and statistical issues. For example, because the demographic characteristics of the eighth-grade public-school students vary widely by jurisdiction, the proportions of students in the various categories of the race/ethnicity, parental education, and type of location variables also varied by jurisdiction. Some of these groups are so small that estimates of statistics describing these groups are too unstable to report. Decisions about minimum sample sizes necessary to report results were made. Chapter 10 documents the major conventions and statistical procedures used in generating the state reports, *Science Report Card*, *Data Compendium*, and the almanacs. The chapter describes the rules, based on effect size and sample size considerations, that were used to establish whether a particular category contained sufficient data for reliable reporting of results for a particular jurisdiction. Chapter 10 also describes the multiple comparison and effect size-based inferential rules that were used for evaluating the statistical and substantive significance of subpopulation comparisons.

To provide information about the generalizability of the results, a variety of information about participation rates was reported for each state and jurisdiction. This included the school participation rates, both in terms of the initially selected samples of schools and in terms of the finally achieved samples, including replacement schools. The student participation rates, the rates of students excluded due to being identified as SD or LEP, and the estimated proportions of assessed students who are classified as SD or LEP were also reported by jurisdiction. These rates are described and reported in Appendix B.

1.14 A SPECIAL SCIENCE ASSESSMENT OF GRADE 4 DoDEA STUDENTS

Many of the features of the special grade 4 DoDEA assessment in science are the same as features of the grade 8 State Assessment in science. Other features vary somewhat from the State Assessment due to the special characteristics of the fourth-grade assessment instruments, the DoDEA schools, and the fact that a fourth-grade assessment of science was not conducted for any other jurisdiction within the nation.

The fourth-grade assessment booklets contain a different proportion of items from each of the fields of science than the eighth-grade assessment booklets. However, the three fields of science are still represented, the numbers and types of blocks remain the same, and the blocks of items are arranged in booklets using the same complex design as for the grade 8 assessment. Chapter 2 contains further information about the difference between the grade 4 and grade 8 instruments.

As described briefly in Chapters 3, 4, and 7, sampling procedures, field administration procedures, and calculation of student weights were the same for the grade 4 DoDEA schools as they were for grade 8 schools. Data entry, professional scoring, and entry into the ETS/NAEP database were similar for all parts of the state and national assessments.

The analysis of the grade 4 DoDEA science data (Ballator, O’Sullivan, & Jerry, 1997a and Ballator, O’Sullivan, & Jerry, 1997b) was somewhat different from the analysis for the grade 8 State Assessment because of the lack of enough grade 4 DoDEA data to scale the items. The resulting analysis is described in Chapter 9. The decisions about reporting results as described in Chapter 10 were applied to the grade 4 DoDEA science assessment in full.

Chapter 2

DEVELOPING THE NAEP 1996 SCIENCE ASSESSMENT INSTRUMENT[1]

Christine O'Sullivan
Educational Testing Service

2.1 OVERVIEW

The science framework for the 1996 National Assessment of Educational Progress (NAEP) was produced under the auspices of the National Assessment Governing Board (NAGB). The consensus process was managed by the Council of Chief State School Officers (CCSSO) who worked with the National Center for Improving Science Education and the American Institutes for Research. Items were developed that were aligned with the specifications described in the framework and were extensively reviewed by specialists in science, measurement, and bias/sensitivity, as well as by government officials and state representatives.

The development of the framework and questions was governed by four major considerations:

- The framework had to be developed through a consensus process involving educators, policy makers, science teachers, representatives of the business community, assessment and curriculum experts, and members of the public.

- The development of the items had to be guided by a Science Instrument Development Committee and receive further reviews by government and state representatives. In addition, the items had to be carefully reviewed for potential bias. (ETS proposal for the administration of the NAEP cooperative agreement 1992.)

- All materials developed at ETS had to be in compliance with specified procedures as described in the ETS Standards of Quality and Fairness (ETS, 1987).

- All NAEP cognitive items and background questions had to be submitted to a federal clearance process as per federal regulations.

This chapter discusses how the specifications and items for the State Assessment in science were developed. It also describes the assessment instrument, the student assessment booklets, and the student, teacher, school, and SD/LEP questionnaires.

[1] Christine O'Sullivan coordinated the development of the science assessment instruments.

Various committees worked on the development of the framework, objectives, and items for the science assessment. The list of committee members and consultants who participated in each aspect of the 1996 development process is provided in Appendix A.

2.2 STEERING COMMITTEE GUIDELINES

The science framework for the NAEP 1996 assessment was developed over a 10-month period between October 1990 and August 1991. The process was initiated with the formation of a Steering Committee that recommended that the framework and ensuing science assessment have the following five characteristics:

- The framework should reflect the best thinking about the knowledge, skills, and competencies needed for a high degree of scientific understanding among all students in the United States. Accordingly it should encompass knowledge and use of organized factual information, relationships among concepts, major ideas unifying the sciences, and thinking and laboratory skills. In addition, the framework should be based on current understandings from research of teaching, learning, and students' performance in science.

- The framework and the assessment should address the nature and practices of knowing in science, as different from other ways of knowing; reflect the quantitative aspects of science as well as the concepts of life, earth, and physical sciences; deal with issues raised by the role of science and technology in society; include practical problem solving in science; take into account the developmental levels of students; and ensure that students with diverse backgrounds are assessed in ways that provide them with equal and fair opportunities to reflect their knowledge and performance.

- Assessment formats should be used that are consistent with the objectives being assessed. A variety of strategies for assessing student performance are advocated, including performance tasks that allow students to manipulate physical objects and draw scientific understandings from the materials before them; constructed-response items that provide insights into students' levels of understanding and ability to communicate in the sciences, as well as their ability to generate, rather than simply recognize information related to scientific concepts and their interconnections; and multiple-choice questions that probe students' conceptual understanding and ability to connect ideas in a scientifically sound way.

- The assessment should contain a broad enough range of questions at different levels of proficiency for identifying three achievement levels for each grade.

- Information on students' demographic and other background characteristics should be collected. Additional information should be collected from students, teachers and administrators about instructional programs and delivery systems, so that their relationships with student achievement can be ascertained and used to inform program and policy decisions.

A Planning Committee was established to identify goals and objectives and to produce the framework. This Planning Committee met monthly from November 1990 through April 1991 and was joined in the first meeting and final meeting by the Steering Committee, which reviewed and reacted to all framework drafts. During this development process, input and reactions were continually sought from a wide range of committee members both within the field of science and external to it. A list of committee members who participated in the developmental process is provided in Appendix A

2.3 FRAMEWORK FOR THE ASSESSMENT

The framework for the 1996 science assessment is represented as a matrix with two dimensions represented by three fields of science (*earth science*, *physical science*, and *life science*) and three elements of knowing and doing science (*conceptual understanding*, *scientific investigation*, and *practical reasoning*). In addition, there are two overarching domains that describe science and nature of science and themes. Figures 2-1 to 2-3, respectively, describe the three fields of science, the elements of knowing and doing science, and the overarching domains.

Figure 2-1
Descriptions of the Three Fields of Science

Earth Science

The earth science component assessed centers on objects and events that are relatively accessible or visible. The concepts and topics covered are solid earth (lithosphere), water (hydrosphere), air (atmosphere), and the earth in space. The solid earth consists of composition; forces that alter its surface; the formation, characteristics, and uses of rocks; the changes and uses of soil; natural resources used by humankind; and natural forces within the earth. Concepts and topics related to water consist of the water cycle; the nature of oceans and their effects on water and climate; and the location of water, its distribution, characteristics, and effect of and influence on human activity. The air is broken down into composition and structure of the atmosphere (including energy transfer); the nature of weather; common weather hazards; and air quality and climate. The earth in space consists of setting of the earth in the solar system; the setting and evolution of the solar system in the universe; tools and technology that are used to gather information about space; apparent daily motions of the sun, the moon, the planets and the stars; rotation of the earth about its axis, the earth's revolution around the sun; and tilt of the earth's axis that produces seasonal variations in the climate.

Physical Science

The physical science component relates to basic knowledge and understanding concerning the structure of the universe as well as the physical principles that operate within it. The major sub-topics probed are matter and its transformations, energy and its transformations, and the motion of things. Matter and its transformations are described by diversity of materials (classification and types and the particulate nature of matter); temperature and states of matter; properties and uses of material (modifying properties, synthesis of materials with new properties); and resource management. Energy and its transformations involve different forms of energy; energy transformations in living systems, natural physical systems, and artificial systems constructed by humans; and energy sources and use, including distribution, energy conversion, and energy costs and depletion. Motion is broken down into an understanding of frames of reference; forces and changes in position and motion; action and reaction; vibrations and waves as motion; general wave behavior; electromagnetic radiation; and the interactions of electromagnetic radiation with matter.

Life Science

The fundamental goal of life science is to attempt to understand and explain the nature and function of living things. The major concepts assessed in life science are change and evolution, cells and their functions, organisms, and ecology. Change and evolution includes diversity of life on earth; genetic variation within a species; theories of adaptation and natural selection; and changes in diversity over time. Cells and their functions consists of information transfer; energy transfer for the construction of proteins; and communication among cells. Organism are described by reproduction, growth and development; life cycles; and functions and interactions of systems within organisms. The topic of ecology centers on the interdependence of life — populations, communities, and ecosystems.

Figure 2-2
Descriptions of Knowing and Doing Science

Conceptual Understanding

Conceptual understanding includes the body of scientific knowledge that students draw upon when conducting a scientific investigation or engaging in practical reasoning. Essential scientific concepts involve a variety of information including facts and events the student learns from science instruction and experiences with the natural environment and scientific concepts, principles, laws, and theories that scientists use to explain and predict observations of the natural world.

Scientific Investigation

Scientific investigation probes students' abilities to use the tools of science, including both cognitive and laboratory tools. Students should be able to acquire new information, plan appropriate investigations, use a variety of scientific tools, and communicate the results of their investigations.

Practical Reasoning

Practical reasoning probes students' ability to use and apply science understanding in new, real-word applications.

Figure 2-3
Descriptions of Overarching Domains

The Nature of Science

The nature of science incorporates the historical development of science and technology, the habits of mind that characterize these fields, and methods of inquiry and problem-solving. It also encompasses the nature of technology that includes issues of design, application of science to real-world problems, and trade-offs or compromises that need to be made.

Themes

Themes are the “big ideas” of science that transcend the various scientific disciplines and enable students to consider problems with global implications. The NAEP science assessment focuses on three themes: systems, models, and patterns of change.

- Systems are complete, predictable cycles, structures or processes occurring in natural phenomena. Students should understand that a system is an artificial construction created to represent, or explain a natural occurrence. Students should be able to identify and define the system boundaries, identify the components and their interrelationships and note the inputs and outputs to the system.

- Models of objects and events in nature are ways to understand complex or abstract phenomena. As such they have limits and involve simplifying assumptions but also possess generalizability and often predictive power. Students need to be able to distinguish the idealized model from the phenomenon itself and to understand the limitations and simplified assumptions that underlie scientific models.

- Patterns of change involve students’ recognition of patterns of similarity and differences, and recognition of how these patterns change over time. In addition, students should have a store of common types of patterns and transfer their understanding of a familiar pattern of change to a new and unfamiliar one.

2.4 DISTRIBUTION OF ASSESSMENT ITEMS

Table 2-1 summarizes the distribution of assessment time across the three *fields of science* — *earth*, *physical*, and *life*. These fields provide the basis for the content area scales. Care was taken to ensure congruence between the proportions used in the assessment (actual) and those indicated in the assessment specifications (recommended).

Table 2-1

Distribution of Assessment Time by Fields of Science, Grade 8

Fields of Science	Recommended	Actual
Earth Science	30%	30%
Physical Science	30%	30%
Life Science	40%	40%

Table 2-2 shows the distribution of assessment time by *knowing and doing science*.

Table 2-2

Distribution of Assessment Time by Knowing and Doing Science, Grade 8

Knowing and Doing Science Elements	Recommended	Actual
Conceptual Understanding	45%	45%
Scientific Investigation	30%	29%
Practical Reasoning	25%	26%

.

A number of items that assess each of the fields of science and each of the ways of *knowing and doing science* also probe *nature of science* and *themes* (*systems, models*, and *patterns of change*). Table 2-3 shows the recommended and actual percentages of assessment time for these two overarching domains.

Table 2-3

Distribution of Assessment Time Devoted to The Nature of Science and Themes, Grade 8

Overarching Domains	Recommended	Actual
Nature of Science	$\geq$15%	21%
Themes	50%	49%

2.5 ASSESSMENT ITEMS, FIELD TEST, AND FINAL FORMS

Items that were closely aligned to the framework were written by teachers from across the country as well as by science assessment specialists on staff at ETS. Several types of items were developed — multiple-choice, short constructed-response, and extended constructed-response. Short constructed-response items (scored with either a 2- or 3-level scoring guide) were used when students needed to respond in a sentence or two. Extended constructed-response items (scored with a 4- or 5-level scoring guide) generally required a paragraph or more. Some items also required diagrams, graphs, or calculations. It was expected that students could adequately answer the short constructed-response questions in about two to three minutes and the extended constructed-response questions in about five minutes. In addition, blocks of items were developed that required the manipulation of equipment (hands-on tasks) and others were developed that assessed each of the three themes: *systems, models,* and *patterns of change*.

Most of the items for the 1996 science assessment were field tested in February and March 1993; however, since the assessment was delayed from 1994 to 1996 an opportunity was afforded for further items to be field tested in February and March 1995. Each of these field tests involved students in many states, the District of Columbia, and three U. S. territories and were intended to try out the cognitive items and hands-on tasks and to give jurisdictions and contractors practice and experience with the proposed materials and tasks. Approximately 500 responses were obtained for each item in each field test.

The field test data were collected, scored, and analyzed in preparation for meetings with the Science Instrument Development Committee. The objectives that guided the review of these items were:

- to determine which items were most suitable for assessing understanding in science in accordance with the framework;
- to determine the need for revisions of questions that lacked clarity, or had ineffective item formats;
- to determine appropriate timing for assessment items.

Committee members, ETS assessment staff, and NAEP staff reviewed the materials. Item analyses (which provided the percentage of correct responses, the r-biserial correlations for multiple-choice and items with a two-level scoring guide, percentages of responses in each category or at each level of the scoring guide, and the r-polyserial for other constructed-response items) were used as a guide in identifying and flagging for further review those test questions that might not measure the intended objective well.

Once the committees had selected the items, they were rechecked for content and measurement concerns and to insure fairness and quality. In addition, a meeting of representatives from state education agencies was convened to review the items chosen for the operational assessment. The federal clearance package containing 13 blocks of cognitive questions was submitted to NCES in August 1993. A further clearance package containing two blocks of items was submitted to NCES in 1995. Throughout the clearance process, revisions were made in accordance with changes required by the government. Upon approval, the 15 blocks (assembled into booklets) and questionnaires were ready for printing in preparation for the assessment.

The following summarizes the series of steps used to create the assessment items for the 1996 State Assessment in science.

1. Item specifications and prototype items were provided in the Science Assessment and Exercise Specifications for the 1996 National Assessment of Educational Progress (National Assessment Governing Board, 1996).

2. The Science Instrument Development Committee provided guidance to NAEP staff about how the objectives could be measured given the realistic constraints of resources and the feasibility of measurement technology. The Committee made recommendations about priorities for the assessment and types of items to be developed.

3. Item writers from both inside and outside ETS were selected based on their knowledge about science education and experience in creating questions according to specifications.

4. The items were reviewed and revised by NAEP/ETS staff and the Instrument Development Committee.

5. Language editing and sensitivity reviews, checking for fairness and quality, were conducted according to ETS quality control procedures.

6. Field test materials were prepared, including the materials necessary to secure clearance by the Office of Management and Budget.

7. The field test was conducted in many states, the District of Columbia, and three territories (see Table 1-1 in Chapter 1).

8. Representatives from State Education Agencies met and reviewed the operational assessment.

9. Based on the field test analyses, items for the 1996 assessment were revised, modified, and re-edited, where necessary. The items once again underwent ETS sensitivity review.

10. The Science Instrument Development Committee approved the selection of items to include in the 1996 assessment.

11. The operational assessment was submitted to NCES for approval.

12. Revisions were made to items in accordance with NCES directives and approval was given.

13. After a final review and check to ensure that each assessment booklet and each block met the overall guidelines for the assessment, the booklets were printed.

2.6 THE ASSESSMENT DESIGN

The 1996 State Assessment in science was made up of 194 cognitive items that were distributed into 15 different sections or blocks. The blocks included four hands-on task blocks, three theme blocks, and eight other paper-and-pencil cognitive blocks. Each block usually contained both multiple-choice and constructed-response items. Each student's booklet consisted of three blocks of cognitive items and students were allowed 30 minutes to complete each block.

Each student assessment booklet also included one section of general background questions (26 items), one section of science background questions (42 items), and one section related to student motivation (5 items). The total administration time for each student for the three cognitive blocks and background items was approximately 103 minutes.

The assembly of science blocks into booklets and their subsequent assignment to sampled students was determined by a complex design with *spiraled* administration. The final cognitive block in each booklet was always one of the hands-on task blocks. The other cognitive blocks were assigned to booklets so that no two theme blocks appeared in the same booklet and every theme block was paired with each of the non-theme paper-and-pencil blocks exactly once. Thirteen booklets contained non-theme paper-and-pencil blocks paired with other non-theme paper-and-pencil blocks. All of the paper-and-pencil blocks appear in the first position and the second position exactly the same number of times. No booklet contained all items and hence there is incomplete data for each assessed student.

Table 2-4 provides the composition of each block of items administered in the 1996 State Assessment program in science. Table 2-5 shows the order of the blocks in each booklet and how the 15 cognitive blocks were arranged across the 37 booklets to achieve the assessment design.

The assessment booklets were then spiraled and bundled. Spiraling involves interweaving the booklets in a systematic sequence so that each booklet appears an appropriate number of times in the sample. The bundles were designed so that each booklet would appear equally often in each position in a bundle. The students within an assessment session were assigned booklets in the order in which the booklets were bundled. Thus, students in an assessment session received different booklets, and only a few students in a session received the same booklet. In most jurisdictions in the State Assessment, up to 30 students were selected from each school, with the aim of providing an initial sample size of approximately 3,000 public school students per jurisdiction per grade, who responded to each item. The nonpublic-school samples differed in size across the jurisdictions, with the number of schools selected proportional to the nonpublic-school enrollment within each jurisdiction. Typically about 20 to 25 nonpublic schools were included for each jurisdiction.

Table 2-4
Block Designations and Assignment
Number of Exercises per Item Type[2]

Block	Type	Total Number of Items	Number of Multiple-Choice Items	Number of Constructed-Response Items	Numbers of Booklets Containing Block
S2BS1	Common Background	26	26	0	37
S2SB1	Science Background	42	42	0	37
S123SB	Motivation Block	5	5	0	37
S2S3	Hands-On Task	6	0	6	10
S2S4	Hands-On Task	10	3	7	9
S23S5	Hands-On Task	8	0	8	9
S12S6B	Hands-On Task	7	0	7	9
S2S7	Theme-Based	12	2	10	8
S23S8B	Theme-Based	10	5	5	8
S12S9B	Theme-Based	13	3	10	8
S2S10	Concept/Problem Solving	16	8	8	6
S2S11	Concept/Problem Solving	16	8	8	6
S23S12	Concept/Problem Solving	16	8	8	6
S23S13	Concept/Problem Solving	16	8	8	6
S12S14B	Concept/Problem Solving	16	7	9	8
S12S15B	Concept/Problem Solving	16	7	9	6
S2S20	Concept/Problem Solving	16	8	8	6
S2S21	Concept/Problem Solving	16	7	9	6

[2] This table documents the number of items, not the percentages of assessment time.

Table 2-5
Booklet Contents for Grade 8 Science Assessment

Booklet Number	Cognitive Block 1	Cognitive Block 2	Hands-On Task Block	Common Background Block	Science Background Block	Science Motivation Block
S201E	S2S7	S2S10	S2S3	S2BS1	S2SB1	S123SB
S202F	S2S7	S2S11	S2S4	S2BS1	S2SB1	S123SB
S203G	S2S7	S23S12	S23S5	S2BS1	S2SB1	S123SB
S204D	S2S7	S23S13	S12S6B	S2BS1	S2SB1	S123SB
S205E	S2S10	S2S11	S2S3	S2BS1	S2SB1	S123SB
S206F	S23S12	S23S8B	S2S4	S2BS1	S2SB1	S123SB
S207G	S2S10	S23S13	S23S5	S2BS1	S2SB1	S123SB
S208D	S2S10	S23S8B	S12S6B	S2BS1	S2SB1	S123SB
S209E	S2S11	S23S12	S2S3	S2BS1	S2SB1	S123SB
S210F	S23S13	S12S14B	S2S4	S2BS1	S2SB1	S123SB
S211G	S2S11	S23S8B	S23S5	S2BS1	S2SB1	S123SB
S212D	S2S11	S12S14B	S12S6B	S2BS1	S2SB1	S123SB
S213E	S23S13	S23S8B	S2S3	S2BS1	S2SB1	S123SB
S214F	S23S8B	S12S15B	S2S4	S2BS1	S2SB1	S123SB
S215G	S23S12	S12S14B	S23S5	S2BS1	S2SB1	S123SB
S216D	S23S12	S12S15B	S12S6B	S2BS1	S2SB1	S123SB
S217E	S23S8B	S12S14B	S2S3	S2BS1	S2SB1	S123SB
S218F	S12S14B	S2S20	S2S4	S2BS1	S2SB1	S123SB
S219G	S23S8B	S2S20	S23S5	S2BS1	S2SB1	S123SB
S220D	S23S13	S2S20	S12S6B	S2BS1	S2SB1	S123SB
S221E	S12S14B	S12S15B	S2S3	S2BS1	S2SB1	S123SB
S222F	S12S15B	S2S21	S2S4	S2BS1	S2SB1	S123SB
S223G	S12S15B	S12S9B	S23S5	S2BS1	S2SB1	S123SB
S224D	S23S8B	S2S21	S12S6B	S2BS1	S2SB1	S123SB
S225E	S2S20	S2S21	S2S3	S2BS1	S2SB1	S123SB
S226F	S2S20	S12S9B	S2S4	S2BS1	S2SB1	S123SB
S227G	S2S20	S2S7	S23S5	S2BS1	S2SB1	S123SB
S228D	S12S14B	S12S9B	S12S6B	S2BS1	S2SB1	S123SB
S229E	S2S21	S12S9B	S2S3	S2BS1	S2SB1	S123SB
S230F	S2S21	S2S7	S2S4	S2BS1	S2SB1	S123SB
S231G	S2S21	S2S10	S23S5	S2BS1	S2SB1	S123SB
S232D	S12S15B	S2S7	S12S6B	S2BS1	S2SB1	S123SB
S233E	S12S9B	S23S13	S2S3	S2BS1	S2SB1	S123SB
S234F	S12S9B	S2S10	S2S4	S2BS1	S2SB1	S123SB
S235G	S12S9B	S2S11	S23S5	S2BS1	S2SB1	S123SB
S236D	S12S9B	S23S12	S12S6B	S2BS1	S2SB1	S123SB
S237E	S12S14B	S2S7	S2S3	S2BS1	S2SB1	S123SB

2.7 BACKGROUND QUESTIONNAIRES

As part of the State Assessment (as well as the national assessment), a series of questionnaires was administered to students, teachers, and school administrators. Similar to the development of the cognitive items, the development of the policy issues and questionnaire items was a consensual process that involved staff work, field testing, and review by external advisory groups. A Background Questionnaire Committee drafted a set of policy issues and made recommendations regarding the design of the items. They were particularly interested in capitalizing on the unique properties of NAEP and not duplicating other surveys (e.g., the National Survey of Public and Private School Teachers and Administrators, the School and Staffing Study, and the National Educational Longitudinal Study). The policy issues, items, and field test results were reviewed by the group of external consultants who identified specific items to be included in the final questionnaires. In addition, the Science Instrument Development Committee and state representatives were consulted on the appropriateness of issues addressed in the questionnaires as they relate to science instruction and performance. The items underwent ETS and NCES review procedures to ensure fairness and quality and were then assembled into questionnaires. The questionnaires were then submitted to the Office of Management and Budget (OMB) for approval.

2.7.1 Student Questionnaires

In addition to three blocks of cognitive items, each booklet in the 1996 State Assessment included three student questionnaires. Two of these were sets of general and science background questionnaires designed to gather contextual information about students, their instructional experiences in science, and their attitudes toward science. The third questionnaire was given to students at the end of each booklet to determine students' motivation in completing the assessment and their familiarity with assessment tasks (see Table 2-5 for placement).

The **student demographics (common background) questionnaire** included questions about race/ethnicity, mother's and father's level of education, types of reading materials in the home, and school attendance.

The **science background questionnaire** included questions that addressed the following.

> *Attitudes Towards Sciences*: Students were asked a series of questions about their attitudes and perceptions about science.
>
> *Time Spent Studying Science*: Students were asked to describe both the amount of instruction they received in science and the time spent on science homework.
>
> *Instructional Practices*: Students were asked to report their instructional experiences related to science in the classroom, including group work, special projects, and writing in response to science. In addition, they were asked about the instructional practices of their science teachers.

The **student motivation questionnaire** asked students how many questions they thought they got right on the NAEP science assessment, how difficult they found it, how hard they tried, how important it was for them to do well, and how often they wrote long answers on tests or assignments for science.

2.7.2 Teacher, School, and SD/LEP Student Questionnaires

To supplement the information on instruction reported by students, the science teachers of the students participating in the State Assessment were asked to complete a questionnaire that addressed teachers' background and general training as well as their science preparation and information concerning science instruction.

The **Teacher Questionnaire, Part I: Background and General Training** included questions about gender, race/ethnicity, years of teaching experience, certification, degrees, major and minor fields of study, course work in education, course work in specific subject areas, amount of in-service training, professional development activities, and availability of resources for their classroom.

The **Teacher Questionnaire, Part II: Science Preparation and Science Instructional Information** included questions on the number and types of science courses taken over the past two years, membership to science organizations, frequency of instructional activities such as asking students to prepare a written science report or an oral science report, emphasis on objectives such as developing science problem-solving skills, methods used to assess student progress in science, and ability level of students in class.

A **School Characteristics and Policies Questionnaire** was given to the principal of each school that participated in the state assessment program. This questionnaire asked about background and characteristics of school principals, length of school day and year, school enrollment, absenteeism, drop-out rates, size and composition of teaching staff, policies about grouping students, curriculum, testing practices and uses, special priorities and school-wide programs, availability of resources, special services, community services, policies for parental involvement, and school-wide problems.

The **SD/LEP Student Questionnaire** was completed by the teachers of those students who were selected to participate in the State Assessment sample and were identified as students with a disability (SD) or were categorized as being of limited English proficiency (LEP). Some of these students were determined by the school to be ineligible to be assessed. In order to be excluded from the assessment, a student must have been identified as SD and must not have been mainstreamed at least 50 percent of the time, or was categorized as LEP. In addition, the school staff would have needed to determine that it was inappropriate to include these students in the assessment. This questionnaire asked about the nature of the student's disability or about the students' native language, and the special programs in which the student participated.

2.8 ASSESSMENT INSTRUMENTS USED IN THE GRADE 4 SCIENCE ASSESSMENT OF DoDEA STUDENTS

The grade 4 student booklets and questionnaires used in the assessment of DoDEA students were the same as those used in the 1996 grade 4 national science assessment. They were developed using the same procedures described in this chapter and for the development of the grade 8 assessment instrument.

There are also similarities between the framework for the grade 4 and grade 8 assessments. The same fields of science, ways of knowing and doing science, and overarching domains were identified for both grades. However, the important specification of the distribution of assessment time varied by grade. Tables 2-6 through 2-8 contain the distribution for fields of science, knowing and doing science, and nature of science and themes, respectively.

Table 2-6

Distribution of Assessment Time by Fields of Science, Grade 4

Fields of Science	Recommended	Actual
Earth Science	33%	33%
Physical Science	33%	33%
Life Science	33%	33%

Table 2-7

Distribution of Assessment Time by Knowing and Doing Science, Grade 4

Knowing and Doing Science Elements	Recommended	Actual
Conceptual Understanding	45%	45%
Scientific Investigation	45%	38%
Practical Reasoning	10%	17%

.

Table 2-8

Distribution of Assessment Time Devoted to The Nature of Science and Themes, Grade 4

Overarching Domains	Recommended	Actual
Nature of Science	≥15%	19%
Themes	33%	53%[1]

[1]Several of the hands-on tasks were classified as themes.

An additional difference between the content of the cognitive items for the two grades was in the area of life science. The grade 4 assessment instrument did not include any items concerning cells and their functions.

The assessment design was parallel for the grade 4 and grade 8 assessments. The same complex spiraled design was used. Student booklets were also laid out in the same way for both grades. Students in grade 4 received only 20 minutes to complete each block of cognitive items, rather than the 30 minutes provided for grade 8 students. As for grade 8, students, teachers, school, and SD/LEP student questionnaires were used at grade 4.

Chapter 3

SAMPLE DESIGN AND SELECTION[1]

John Burke and James L. Green
Westat, Inc.

3.1 OVERVIEW

The 1996 State Assessment program in mathematics included assessments of fourth- and eighth-grade students in public- and nonpublic-schools. For the eighth-grade, the samples selected for both the mathematics and science assessment were selected as part of the same process. Some schools that were selected for participation in the eighth-grade sample provided both students that were assessed in mathematics and students that were assessed in science. (This was also true for the DDESS and DoDDS sample.) A representative sample of public- and nonpublic-school students was drawn in each participating jurisdiction. Each sample was designed to produce aggregate estimates as well as estimates for various subpopulations of interest with approximately equal precision for the participating jurisdictions. The sample for the fourth- and eighth-grade public-school assessments in each jurisdiction consisted of about 3,150 assessments (before attrition) in each subject from about 100 public schools in each case. The target for nonpublic-school assessments varied by jurisdiction and was proportional to their representation in the jurisdiction.

The target population for the 1996 State Assessment program included students in public and nonpublic schools who were enrolled in the fourth and eighth grade at the time of assessment. The sampling frame included public and nonpublic schools having the relevant grade in each jurisdiction. The samples were selected based on a two-stage sample design; selection of schools within participating jurisdictions, and selection of students within schools. The first-stage samples of schools were selected with probability proportional to a measure of size based on the estimated grade-specific enrollment in the schools. Special procedures were used for jurisdictions with many small schools, and for jurisdictions having small numbers of grade-eligible schools.

Stratification variables were added to the sampling frame prior to sample selection. Public schools were stratified by urbanization and minority class and nonpublic schools were stratified by metro area status and school type. The urbanization strata were defined in terms of large or mid-size central city, urban fringe of large or mid-size city, large town, small town, and rural areas. Within urbanization strata, public schools were further stratified explicitly on the basis of minority enrollment in those jurisdictions with substantial Black or Hispanic student population. Minority enrollment was defined as the total percent of Black and Hispanic students enrolled in a school. Within minority strata, public schools were sorted by median household income of the ZIP code area where the school was located. Metro area status was determined by U.S. Bureau of Census definitions as of June 30, 1993. School type was a dichotomous variable (Catholic or other nonpublic). Within school type, nonpublic schools were sorted by estimated grade enrollment.

---

[1] John Burke was responsible for overseeing all sampling activities; James Green carried out most of these activities.

From the stratified frame of public schools within each jurisdiction, a systematic random sample of about 100 grade-eligible schools was drawn with probability proportional to a measure of size based on the estimated grade-specific enrollment of the school. Each selected school provided a list of eligible enrolled students, from which a systematic sample of students was drawn. One or more sessions of 30 students were sampled within each school. The number of sessions selected depended on the school's estimated grade-specific enrollment, though the overwhelming majority of schools at grade 4 were allocated a single session.

One fourth of the selected public schools were designated at random to be monitored during the assessment field period so that reliable comparisons could be made between sessions administered with and without monitoring. This was done in all jurisdictions that participated in the 1994 Trial State Assessments. One half of the selected public schools were designated to be monitored in jurisdictions that did not participate in the 1994 Trial State Assessments.

Approximately 3,150 public-school students were targeted for selection for a given grade and subject in a given jurisdiction. On average, 109 public schools and 20 nonpublic schools were selected for fourth grade in each jurisdiction and 105 public schools and 31 nonpublic schools were selected for eighth grade in each jurisdiction. The maximum number of public and nonpublic schools sampled in a jurisdiction were 139 and 44, respectively, for fourth grade. The minimum number of public and nonpublic schools sampled in a jurisdiction were 22 and 10, respectively, for fourth grade. The maximum number of public and nonpublic schools sampled for eighth grade were 159 and 68, respectively. The minimum number of public and nonpublic schools sampled in a jurisdiction were 6 and 10, respectively, for eighth grade.

The 1996 State Assessment was preceded in 1995 by a field test. The principal goals of the field test were: 1) to test new items contemplated for 1996, and 2) to test procedures contemplated for 1996. Schools that participated in the field test were given a chance of selection in the 1996 assessment. Section 3.2 documents the procedures used to select the schools for the field test.

Section 3.3 describes the construction of the sampling frames, including the sources of school data, missing data problems, and definition of in-scope schools. Section 3.4 includes a description of the various steps in stratification of schools within participating jurisdictions. School sample selection procedures (including new and substitute schools) are described in Section 3.5. Section 3.5.5 includes information about the selection of schools for application of the two sets of inclusion rules (S1 and S2 subsamples) used in the State Assessment. Section 3.6 includes the steps involved in selection of students within participating schools.

3.2 SAMPLE SELECTION FOR THE 1995 FIELD TEST

The 1995 field test for the State Assessment program was conducted together with the field test for the national portion of the assessment. In these field tests, assessments were piloted in: mathematics, science, and the arts (dance, music, theater, and visual arts). All jurisdictions were included in the field test except Alaska, Delaware, the District of Columbia, Hawaii, Rhode Island and Wyoming, which were excluded due to the heavy burden placed on these small population jurisdictions by the main assessment. The field test was conducted for grades 4, 8, and 12. Groups of three schools were identified as described in Section 3.2.2, with one school from each group to be included in the test. This allowed state participation in the selection of the test schools and also facilitated replacement of schools that declined to participate in the assessment. Sampling weights were not computed for the field test samples.

3.2.1 Primary Sampling Units

The field test primary sampling units (PSU) sampling frame was derived from the national list of U.S. counties. The frame was stratified by state and metro area status. Two hundred and fifteen PSUs were selected from the resulting field test frame. Twenty PSUs were selected with certainty and 195 noncertainty PSUs were selected — one per noncertainty stratum. The PSUs were selected systematically and with probability proportional to the 1990 PSU population. Counties that were noncertainty selections for the Third International Mathematics and Science Study (TIMSS), the 1996 NAEP national assessment, and the 1996 NAEP trend samples were excluded from the sampling frame. The number of counties selected per jurisdiction ranged from 2 to 10.

3.2.2 Selection of Schools and Students

Public and nonpublic schools with fourth-, eighth-, or twelfth-grade students were in-scope for the field test assessment. Schools with fewer than 40 students were eliminated from the sampling frame to avoid the relatively high per student cost of conducting assessments in small schools. Schools selected as originals or substitutes for TIMSS were also eliminated from the frame.

Across all three grades from the resulting sampling frame, 1,285 groups of three schools were selected. The first member of each group of schools was selected systematically and with probability proportional to grade enrollment. The twelfth-grade sample was drawn first followed by the eighth- and fourth- grade samples. The selected twelfth-grade schools were removed from the frame before drawing the eighth-grade sample. The selected twelfth- and eighth-grade schools were removed from the frame before drawing the fourth-grade sample. In this way, no school was selected for more than one grade.

The second member of each group of three schools was selected within the same district as the first member and in such a way that the "distance" (described in Section 3.5.6) from the preliminary selection, based on percent of Black students, percent of Hispanic students, grade enrollment, and percent of students living below poverty, was the smallest across all schools remaining after the fourth-, eighth-, and twelfth-grade sampling. The third member was selected similarly, but in a different district in the same PSU as the first and second members. In some cases, a third member was not available for each group of schools.

3.2.3 Assignment to Sessions for Different Subjects

Up to six different session types were assigned in a given jurisdiction. The particular number of session types varied by grade. Table 3-1 gives the overall number of schools selected for each grade and session type.

Table 3-1

Number of Schools Selected for the Field Test for Each Grade and Session Type

Session Type[1]	Grade 4	Grade 8	Grade 12
Mathematics/Science	75	80	90
Mathematics Trend	75	75	75
Visual Arts/Music	75	70	100
Theater and Dance	85	85	120
Spanish/Bilingual	120	120	0
SD Accommodations	20	20	0
Total	**450**	**450**	**385**

[1]The mathematics and science sessions were sessions where items selected for the 1996 State Assessment program were administered. The mathematics trend sessions were sessions where booklets from the 1996 Trial State Assessment program in mathematics were administered. The results from students included in these sessions were used to verify that the 1996 and 1992 assessments could be placed on the same scale. The two types of arts sessions were selected for every grade. They were administered at grades 4 and 8, although the grade 12 sessions were administered the arts field test in 1997. The Spanish/bilingual and SD accommodations sessions were administered using special booklets to determine whether SD/LEP students could participate in the 1996 assessment with special accommodations.

The number of sessions assigned to an individual school depended on the size of the school and the subject(s) that school was assigned.

3.3 TARGET POPULATION AND SAMPLING FRAME FOR THE 1996 ASSESSMENT

3.3.1 Target Population

The target population for the 1996 State Assessment included students in public and nonpublic schools who were enrolled in the fourth or eighth grade. Nonpublic schools included Catholic schools, other religious schools, private schools, DDESS, and Bureau of Indian Affairs (BIA) schools. Special education schools were not included. Both S1, based on the old inclusion rules, and S2, based on the new inclusion rules, shared this target population (see Chapter 4).

3.3.2 Sampling Frame

In order to draw the school samples for the 1996 State Assessment, it was necessary to obtain a comprehensive list of public and nonpublic schools in each jurisdiction. For each school, useful information for stratification purposes, reliable information about grade span and enrollment, and accurate information for identifying the school to the state coordinator (district membership, name, address) were required.

Based on the experience with the 1992 and 1994 Trial State Assessments, and national assessments from 1984 to 1994, the file made available by Quality Education Data, Inc. (QED) was elected as the sampling frame. The National Center for Education Statistics' Common Core of Data (CCD) school file was used to check the completeness of the QED file.

The QED list covers all U.S. states and jurisdictions except Puerto Rico. The version of the QED file used was released in late 1994, in time for selection of the school sample in early 1995. The file was missing racial/ethnic minority enrollment and urbanization data for a sizable minority of schools (due to the inability of QED to match these schools with the corresponding CCD file). Considerable efforts were undertaken to obtain these variables for all schools in jurisdictions where these variables were to be used for stratification. These efforts are described in the next section.

A new addition for 1996 was the joint use of QED and National Center for Education Statistics' Private School Universe Survey (PSS) lists of nonpublic schools. These two sources were combined, eliminating duplicates as necessary and increasing coverage throughout the combined frame. When a given school was found on both lists, the PSS data were given priority.

Tables 3-2 and 3-3 show the distribution of fourth- and eighth-grade schools and enrollment within schools as reported in the combined frame. Grade-specific enrollment was estimated for each school as the quotient of total school enrollment and the number of grades in the school.

Table 3-2

Distribution of Fourth-Grade Schools and Enrollment in Combined Frame

	Public Schools		Nonpublic Schools	
Jurisdiction	**Total Schools**	**Total Enrollment**	**Total Schools**	**Total Enrollment**
DoDEA/DDESS	39	3,118	N/A	N/A
DoDEA/DoDDS	113	7,291	N/A	N/A
Total	**152**	**10,415**	**N/A**	**N/A**

Table 3-3

Distribution of Eighth-Grade Schools and Enrollment In Combined Frame

	Public Schools		Nonpublic Schools	
Jurisdiction	**Total Schools**	**Total Enrollment**	**Total Schools**	**Total Enrollment**
Alabama	484	56,995	245	5,363
Alaska	256	9,240	59	481
Arizona	328	54,351	227	4,210
Arkansas	344	35,074	110	1,846
California	1,642	379,030	2,023	47,939
Colorado	325	46,695	224	3,795
Connecticut	207	34,383	248	5,828
Delaware	29	7,751	101	2,097
District of Columbia	32	4,808	46	1,435
DoDEA/DDESS	12	1,517	N/A	N/A
DoDEA/DoDDS	66	5,353	N/A	N/A
Florida	466	152,838	839	19,767
Georgia	398	97,029	385	8,297
Guam	6	2,199	12	498

(continued)

Table 3-3 (continued)
Distribution of Eighth-Grade Schools and Enrollment In Combined Frame

	Public Schools		Nonpublic Schools	
Jurisdiction	**Total Schools**	**Total Enrollment**	**Total Schools**	**Total Enrollment**
Hawaii	52	12,845	84	3,341
Indiana	437	76,101	558	9.073
Iowa	409	38,331	194	4,461
Kentucky	357	51,275	238	6,293
Louisiana	431	59,102	352	13,767
Maine	235	16,134	98	1,077
Maryland	229	57,586	383	9,942
Massachusetts	383	61,789	407	10,656
Michigan	737	120,422	819	16,577
Minnesota	424	59,224	361	7,447
Mississippi	301	39,570	143	4,076
Missouri	633	63,768	441	10,375
Montana	321	12,800	81	834
Nebraska	577	22,137	173	3,502
Nevada	95	18,626	44	840
New Hampshire	132	14,600	78	1,228
New Jersey	664	84,346	660	18,516
New Mexico	152	24,249	148	2,387
New York	1,013	187,305	1,368	42,412
North Carolina	526	89,074	377	6,856
North Dakota	237	9,065	54	743
Oregon	343	39,630	195	2,808
Rhode Island	52	10,286	77	2,163
South Carolina	252	51,010	206	3,679
Tennessee	533	66,684	325	7,044
Texas	1,488	271,798	680	16,095
Utah	142	36,877	54	913
Vermont	126	7,413	52	575
Virginia	336	79,009	362	7,124
Washington	425	70,998	345	6,430
West Virginia	206	24,448	126	1,214
Wisconsin	513	61,628	778	13,729
Wyoming	96	7,971	31	195
Total	**21,740**	**3,243,013**	**18,452**	**423,591**

3.4 STRATIFICATION

3.4.1 Stratification Variables

The stratification used for sample selection varied by school type (public or nonpublic). Stratification of public schools involved four primary dimensions whereas the stratification of nonpublic schools involved three primary dimensions. Public schools were stratified hierarchically by small or large district status, school size class, urbanization classification and minority classification. Nonpublic schools were stratified by size class, metro area status and school type (Catholic or other nonpublic). Public schools were further stratified implicitly by median household income of the ZIP code area where the school was located (i.e., sorted in ascending or descending order) and nonpublic schools were further stratified implicitly by estimated grade enrollment in order to provide some control of these variables. The DDESS schools, the DoDDS schools (except fourth grade), and Guam samples are not included in these tables as all schools in these jurisdictions were sampled with certainty, thereby requiring no stratification. The DoDDS fourth-grade sample was sorted by Department of Defense Education Activity (DoDEA) area (Europe, Pacific, etc.), DoDEA district (Brussels, Heidelberg, Italy, etc.), and estimated fourth-grade enrollment prior to sample selection.

3.4.2 Missing Stratification Variables

As stated earlier, the sampling frame for the 1996 State Assessment was the most recent version of the QED file available combined with the 1993 PSS list of nonpublic schools. The CCD file was used to extract information on urbanization ("type of location") for public schools where this information was missing on the QED file. Any public schools with missing values remaining in urbanization or minority enrollment data were imputed.

Schools with missing values in urbanization data were assigned the urbanization of other school records within the same state, county, and city when urbanization did not vary within the given city. Any schools still missing urbanization were assigned the modal value of urbanization within their city. Any remaining missing values were assigned individually based on city using U.S. Bureau of Census publications.

Schools with missing values in minority enrollment data were assigned the average minority enrollment within their school district. Any schools still missing minority enrollment data were assigned values individually using ZIP code and U.S. Bureau of Census data. The minority data were extracted only for those schools in jurisdictions in which minority stratification was performed.

Metro area status was assigned to each nonpublic school based on U.S. Bureau of Census definitions as of June 30, 1993, based on Federal Information Processing Standard (FIPS) county code, and was found for all schools in the sampling frame. The Catholic school flag was assigned to each nonpublic school based on the QED or PSS school type and was found for all schools in the sampling frame.

Median household income was assigned to every school in the sampling frame by merging on ZIP code with a file from Donnelly Marketing Information Services. Any schools still missing median household income were assigned the mean value of median household income for the three-digit ZIP code prefix or county within which they were located.

3.4.3 Urbanization Classification

Urbanization classification was created based on the NCES type of location variable. The type of location variable contains at most seven levels:

1. *Large Central City*: A central city of a Metropolitan Statistical Area (MSA) with a population greater than or equal to 400,000, or a population density greater than or equal to 6,000 persons per square mile;

2. *Mid-size Central City*: A central city of an MSA but not designated as a large central city;

3. *Urban Fringe of Large City*: A place within an MSA of a large central city and defined as urban by the U.S. Bureau of Census;

4. *Urban Fringe of Mid-Size City*: A place within an MSA of a mid-size central city and defined as urban by the U.S. Bureau of Census;

5. *Large Town*: A place not within an MSA, but with a population greater than or equal to 25,000 and defined as urban by the U.S. Bureau of Census;

6. *Small Town*: A place not within an MSA, with a population less than 25,000, but greater than 2,499 and defined as urban by U.S. Bureau of Census; and

7. *Rural*: A place with a population of less than 2,500 and defined as rural by the U.S. Bureau of Census.

Urbanization classification was created by collapsing type of location categories as necessary and according to specific rules until each urbanization stratum included a minimum of 10 percent of eligible students in the participating jurisdiction. Table F-1 in Appendix F provides the urbanization classifications used within each jurisdiction for grade 8.

3.4.4 Minority Classification

Minority classification was created within urbanization strata and was based on a school's percentages of Black and Hispanic students. Three different minority classification schemes were used and are described as follows:

- *Case 1:* Urbanization strata with less than 10 percent Black students and 7 percent Hispanic students were not stratified by minority enrollment (Level 0);

- *Case 2:* Urbanization strata with greater than or equal to 10 percent Black students or 7 percent Hispanic students, but not more than 20 percent of each, were stratified by ordering percent minority enrollment (Black plus Hispanic) within the urbanization classes and dividing the schools into three groups with about equal numbers of students per minority classification (Levels 1, 2, and 3); and

- *Case 3:* In urbanization strata with greater than 20 percent of both Black and Hispanic students, minority strata were formed with the objective of providing equal strata with emphasis on the minority group (Black or Hispanic) with the higher concentration. The stratification was performed as follows. The minority group with the higher percentage gave the primary stratification variable; the remaining group gave the secondary stratification variable. Within urbanization class, the schools were sorted based on the primary stratification variable and divided into two groups of schools containing approximately equal numbers of students based on estimated grade enrollment. Within each of these two groups, the schools were sorted by the secondary stratification variable and subdivided into two subgroups of schools containing approximately equal numbers of students. As a result, within urbanization strata there were four minority classifications (e.g., low Black/low Hispanic, low Black/high Hispanic, high Black/low Hispanic, and high Black/high Hispanic (Levels 4, 5, 6, and 7).

The minority groups and classifications were formed solely for the purpose of creating efficient stratification design at this stage of sampling. These classifications are not directly used in analysis and reporting of the data, but will act to reduce sampling errors. Table F-1 in Appendix F provide information on minority stratification for the participating jurisdictions for grade 8.

3.4.5 Median Household Income

Prior to the selection of the school samples, the public schools were sorted by their four stratification variables (small or large district status, school class size, urbanization classification, and minority classification) in an order such that changes occur on only one variable at a time (also known as a serpentine order.) This is accomplished by alternating between ascending and descending sort order on each variable successively through the sort hierarchy. Within this sorted list, the schools were sorted, in serpentine order, by the median household income. This final stage of sorting resulted in implicit stratification of median household income. The data on median household income was related to the ZIP code area in which the school is located. The data were derived from the 1990 Census and are obtained from Donnelly Marketing Information Services.

3.4.6 Metro Area Status

All schools in the sampling frame were assigned a metro area status based on their FIPS county code and Office of Management and Budget (OMB) Metropolitan Area Definitions as of June 30, 1993. This field indicated if a school was located within a metropolitan area or not. Table F-2 in Appendix F provides information on metro area status stratification for the participating jurisdictions for grade 8.

3.4.7 School Type

All nonpublic schools were assigned a school type (Catholic or other nonpublic) based on their QED or PSS school-type variable. Table F-2 in Appendix F provides information on school-type stratification for the participating jurisdictions for grade 8.

3.5 SCHOOL SAMPLE SELECTION

3.5.1 Measure of Size and Sample Selection

Each grade-eligible school was assigned an estimated grade enrollment by dividing its total student enrollment by its number of grades. Each school was then assigned a measure of size based on the following function of estimated grade enrollment (EGE). Tables 3-4 and 3-5 provide the estimated grade enrollment and measure of size for both subject areas for grades 4 and 8.

Table 3-4
Estimated Grade Enrollment and Measure of Size, Grade 4

Estimated Grade Enrollment	Measure of Size
EGE < 10	15
10 ≤ EGE < 20	1.5 * EGE
20 ≤ EGE < 33	30
33 ≤ EGE	EGE

Table 3-5
Estimated Grade Enrollment and Measure of Size, Grade 8

Estimated Grade Enrollment	Measure of Size
EGE < 10	30
10 ≤ EGE < 20	3 * EGE
20 ≤ EGE < 65	60
65 ≤ EGE	EGE

Schools were designated as being in “small” or “large” districts and were assigned to one of two size classes, as shown in Tables 3-4 and 3-5. A large district was defined as a district containing 20 percent or more of a jurisdiction’s student population. All other districts were considered small. Schools were assigned to the large size class if their estimated grade enrollment was greater than 19. Otherwise schools were assigned to the small size class.

A sample of schools was then selected for each jurisdiction with probability proportional to each school’s measure of size. The sampling frame of schools was sorted in systematic order prior to sample selection, as follows:

- Public schools
  - ⇒ Small or large district status,
  - ⇒ Size class,
  - ⇒ Urbanization stratum,
  - ⇒ Minority stratum, and
  - ⇒ Median household income.

- Nonpublic schools
  - ⇒ Size class,
  - ⇒ Metro area status,
  - ⇒ Catholic/non Catholic, and
  - ⇒ Estimated grade enrollment.

Sorting the sampling frame in a specific order prior to systematic sample selection ensures that the sampled units represent a variety of population subgroups.

3.5.2 Control of Overlap of School Samples for National Educational Studies

The issue of school sample overlap has been relevant in all rounds of NAEP in recent years. To avoid undue burden on individual schools, NAEP developed a policy for 1996 of avoiding overlap between national and state samples. This was to be achieved without unduly distorting the resulting samples by introducing bias or substantial variance. The procedure used was an extension of the method proposed by Keyfitz (1951). The general approach is given in *The NAEP 1994 Technical Report* (Allen, Kline, & Zelenak, 1996). Counts of school selection for both state and national NAEP are found in Table 3-6.

Table 3-6

Number of Schools Selected for Both State and National NAEP, by Grade and School Type

State NAEP		National NAEP Grade					
Grade	**School Type**	**4 Main**	**4 Trend**	**8 Main**	**8 Trend**	**12 Main**	**12 Trend**
4	Public	10	29	9	4	4	0
4	Nonpublic	0	2	17	1	11	5
8	Public	8	4	53	101	26	4
8	Nonpublic	23	5	5	5	22	4

3.5.3 Selection of Schools in Small Jurisdictions

All schools in jurisdictions with small numbers of public schools for both subject areas were selected. The jurisdictions and grades are shown in Table 3-7.

Table 3-7
Jurisdictions Where All Public Schools were Selected, by Grade and School Type

	Public		Nonpublic	
Jurisdiction	Grade 4	Grade 8	Grade 4	Grade 8
Delaware	*	*		
District of Columbia	*	*		*
DoDEA/DDESS	*	*	N/A	N/A
DoDEA/DoDDS		*	N/A	N/A
Guam	*	*	*	*
Hawaii		*		
Rhode Island		*		

3.5.4 New School Selection

A sample of new schools was drawn to properly reflect additions to the target population occurring after the sampling frame building information was created.

A district-level file was constructed from the combined QED and PSS school-level files. The district-level file was divided into a small districts file, consisting of those districts in which there were at most three schools on the aggregate frame and no more than one fourth-, one eighth-, and one twelfth-grade school. The remainder of districts were denoted as “medium and large” districts.

A sample of medium and large districts was drawn in each jurisdiction. All districts were selected in Delaware, the District of Columbia, Hawaii, and Rhode Island. The remaining jurisdictions in the file of medium and large districts (eligible for sampling) were divided into two files within each jurisdiction. Two districts were selected per jurisdiction with equal probability among the smaller districts with combined enrollment of less than or equal to 20 percent of the state enrollment. From the rest of the file, eight districts were selected per jurisdiction with probability proportional to enrollment. The breakdown given above applied to all jurisdictions except Alaska and Nevada, where four and seven districts were selected with equal probability and six and three districts were selected with probability proportional to enrollment, respectively. The 10 selected districts in each jurisdiction were then sent a listing of all their schools that appeared on the file, and were asked to provide information about the new schools not included in the file. These listings, provided by selected districts, were used as sampling frames for selection of new schools.

The eligibility of a school was determined based on the grade span. A school was also classified as “new” if a change of grade span was such that the school status changed from ineligible to eligible. The average grade enrollment for these schools was set to the average grade enrollment before the grade-span change. The schools found eligible for sampling due to the grade-span change were added to the new school selection frame.

The probability of selecting a school was minimum $\left[\dfrac{\text{sampling rate} \;*\; \text{measure of size}}{P(\text{district})}, 1\right]$,

where P(district) was the probability of selection of a district and the sampling rate was the rate used for the particular jurisdiction in the selection of the original sample of schools.

In each jurisdiction, the sampling rate used for the main sample of grade-eligible schools was used to select the new schools. Additionally, all new eligible schools coming from small districts (those with at most one grade 4 and one grade 8 school) that had a school selected in the regular sample for the fourth grade were included in the sample with certainty. In the 1996 State Assessment, there were no such schools.

Table 3-8 shows the number of new schools coming from the medium and large and small districts for the eighth-grade samples of the 1996 State Assessment in science. There were no new schools for the DoDEA sample.

Table 3-8
Distribution of New Schools Coming from "Medium" or "Large" and "Small" Districts in the Eighth-Grade Sample

	Number of New Schools	
Jurisdiction	**"Medium" or "Large" Districts**	**"Small" Districts**
Alabama	0	0
Alaska	0	0
Arizona	1	0
Arkansas	1	0
California	0	0
Colorado	1	0
Connecticut	0	0
Delaware	2	0
DoDEA/DDESS	0	0
DoDEA/DoDDS	0	0
District of Columbia	2	0
Florida	4	0
Georgia	2	0
Guam	0	0
Hawaii	1	0
Indiana	1	0
Iowa	0	0
Kentucky	0	0
Louisiana	3	0
Maine	0	0
Maryland	1	0
Massachusetts	1	0
Michigan	0	0
Minnesota	0	0
Mississippi	2	0
Missouri	1	0
Montana	1	0
Nebraska	3	0
Nevada	2	0
New Hampshire	0	0

(continued)

Table 3-8 (continued)
Distribution of New Schools Coming from "Medium" or "Large" and "Small" Districts in the Eighth-Grade Sample

	Number of New Schools	
Jurisdiction	**"Medium" or "Large" Districts**	**"Small" Districts**
New Jersey	0	0
New Mexico	0	0
New York	0	0
North Carolina	1	0
North Dakota	0	0
Oregon	2	0
Rhode Island	0	0
South Carolina	1	0
Tennessee	2	0
Texas	1	0
Utah	3	0
Vermont	2	0
Virginia	0	0
Washington	1	0
West Virginia	0	0
Wisconsin	3	0
Wyoming	4	0
Total	**55**	**0**

3.5.5 Assigning Subject, Sample Type, and Monitor Status

Subject assignment rules varied by grade. All fourth grade schools were assigned to participate in mathematics assessments except for the DDESS and DoDDS samples where the rules for subject assignment at eighth grade were followed. All eighth-grade schools with 20 or more students were assigned to participate in both mathematics and science assessments. Schools with less than 20 students were assigned one subject selected at random.

The 1996 State Assessment used two different sets of inclusion rules (see Chapter 4) for different sets of schools (S1 and S2 subsamples). A sample type variable was created to reflect which set of rules to use within a given school. The sampled schools were sorted by stratum (public and nonpublic) and subject (both mathematics and science, mathematics only, and science only) and then randomly assigned sample type within the sorted list. The sets of inclusion rules are described in Chapter 4.

Jurisdictions received 25 or 50 percent monitoring of sessions depending on previous participation in the state assessments. All jurisdictions received 25 percent monitoring except Alaska, Nevada, Vermont, and Washington, where 50 percent monitoring was used. The sampled schools were sorted by stratum, subject, and sample type, and then assigned the two levels of monitoring at random.

3.5.6 School Substitution and Retrofitting

A substitute school was assigned to each sampled school (to the extent possible) prior to the field period through an automated substitute selection mechanism that used distance measures as the matching criterion. Schools were also required to be of the same type (i.e., public, nonpublic, BIA, and DoDEA schools were only allowed to substitute for each other), and substitutes for nonpublic, BIA, and DoDEA schools were required to come from within the same district. Public-school substitutes were required to come from different districts. Two passes were made at the substitution, with the second pass raising the maximum distance measure allowed and removing the different district assignment for public schools. This strategy was motivated from the fact that most public-school nonresponse is really at the district level.

A distance measure was used in each pass and was calculated between each sampled school and each potential substitute. The distance measure was equal to the sum of four squared standardized differences. The differences were calculated between the sampled and potential substitute school's estimated grade enrollment, median household income, percent Black enrollment and percent Hispanic enrollment. Each difference was squared and standardized to the population standard deviation of the component variable (e.g., estimated grade enrollment) across all grade-eligible schools and jurisdictions. The potential substitutes were then assigned to sampled schools by order of increasing distance measure. An acceptance limit was put on the distance measure of .60 for the first pass. A given potential substitute was assigned to one and only one sampled school. Some sampled schools did not receive assigned substitutes (at least in the first pass) because the number of potential substitutes was less than the number of sampled schools or the distance measure for all remaining potential substitutes from different districts was greater than .60.

In the second pass, the different district constraint for public schools was lifted and the maximum distance allowed was raised to .75. This generally brought in a small number of additional assigned substitutes. Although the selected cut-off points of .60 and .75 on the distance measure were somewhat arbitrary, they had been decided upon for the 1994 Trial State Assessment by a group of statisticians reviewing a large number of listings beforehand and finding a consensus on the distance measures at which substitutes began to appear unacceptable.

Jurisdictions that did not receive substitutes for all selected schools were allowed to retrofit unused substitutes after part of the field period elapsed. Substitutes that were assigned to cooperating or ineligible original selections were free to be assigned to other original selections that did not receive substitutes. These free substitutes were put back into the substitute selection mechanism described above and allowed to pair up with other original selections.

Cooperating original selections were also allowed to serve as "double session" substitutes for other pending or refusing schools and were put through the substitute selection mechanism after retrofitting unused regular substitutes. Double session substitutes are particularly helpful to small jurisdictions where all or most schools are taken as original selections, thereby leaving no or few schools available as substitutes.

Tables F-3 and F-4 in Appendix F include information about the number of substitutes provided in each jurisdiction. Of the 47 participating jurisdictions, 41 were provided with at least one substitute at grade 8. Among jurisdictions receiving no substitutes, the majority had 100 percent participation from the original sample.

Tables F-5 through F-7 in Appendix F show the number of schools in the fourth- and eighth-grade science samples, together with school response rates observed within participating jurisdictions. The tables also show the number of substitutes in each jurisdiction that were associated with a nonparticipating original school selection, and the number of those that participated.

3.6 STUDENT SAMPLE SELECTION

3.6.1 Student Sampling and Participation

Schools initially sent a complete list of students to a central location in November 1995. They were not asked to list students in any particular order, but were asked to implement checks to ensure that all grade-eligible students were listed. Based on the total number of students on this list, the "Student Listing Form," sample line numbers were generated for student sample selection. To generate these line numbers, the sampler entered the number of students on the form and the number of sessions into a calculator or personal computer that had been programmed with the sampling algorithm. The program generated a random start that was used to systematically select the student line numbers (30 per session). To compensate for new enrollees not on the Student Listing Form, extra line numbers were generated for a supplemental sample of new students.

After the student sample was selected, the administrator at each school identified students who were incapable of taking the assessment either because they were identified as students with disabilities (SD) or because they were classified as being of limited English proficiency (LEP). Two different sets of inclusion rules were used: a set used in previous assessments and a new set that was meant to clarify the inclusion rules used in NAEP and to provide wider inclusion of SD and LEP students (see Olson & Goldstein, 1997).

When the assessment was conducted in a given school, a count was made of the number of nonexcluded students who did not attend the session. If this number exceeded three students, the school was instructed to conduct a makeup session, to which all students who were absent from the initial session were invited.

Tables F-8 through F-10 in Appendix F provide the distribution of the student samples and response rates by grade, school type, and jurisdiction.

3.6.2 The Reduced Sample Option

All jurisdictions were given the option to reduce the expected student sample size in order to reduce testing burden and the number of multiple-testing sessions for participating schools. If jurisdictions chose to exercise this option, the estimates obtained from the assessment were more variable than they otherwise would have been. In general, jurisdictions could reduce student sample sizes by adjusting the number of sessions with participating schools subject to the following constraints:

- The minimum number of sessions per school had to be equal to 1;

- The maximum number of sessions per school had to be equal to 2 at fourth grade and 3 at eighth grade;

- The expected student sample size from the reduced sample was greater than or equal to half of the original student sample size.

Table 3-9 shows the jurisdictions that exercised the reduced sample option at each grade for both subject areas.

Table 3-9
Jurisdictions Exercising the Reduced Sample Option, By Grade

Jurisdiction	Grade 4	Grade 8
Alaska		*
Delaware	*	*
Guam	*	
Hawaii		*
Rhode Island		*

Chapter 4

STATE AND SCHOOL COOPERATION AND FIELD ADMINISTRATION[1]

Nancy W. Caldwell
Westat, Inc.

4.1 OVERVIEW

By volunteering to participate in the State Assessment and in the field test that preceded it, each jurisdiction assumed responsibility for securing the cooperation of the schools sampled by NAEP. The participating jurisdictions were responsible for the actual administration of the 1996 State Assessment at the school level. The 1995 field test, however, operated within the framework of the national (rather than state) model. Therefore, for the field test, NAEP field staff were responsible for securing cooperation for, scheduling, and conducting the assessments. This chapter describes state and school cooperation and field administration procedures for both the 1995 field test and the 1996 assessment program. Section 4.2 presents information on the field test, while Section 4.3 focuses on the 1996 State Assessment.

4.2 THE FIELD TEST

4.2.1 Conduct of the Field Test

In preparation for the 1996 state and national assessment programs, a field test of the forms, procedures, and booklet items was held in late January through early March 1995. In this field test, assessments were piloted in: mathematics, science, and the arts (dance, music, theater, and visual arts). In an effort to increase the participation of limited English proficient (LEP) students and students with disabilities (SD), the mathematics field test included bilingual and Spanish-language versions of three test booklets, newly developed Braille and large-print booklets, and the provision of additional testing accommodations for students with disabilities and students with limited English proficiency. Results for the field testing of the Spanish-language mathematics assessment, Braille and large-print booklets, and special testing accommodations are contained in a separate report prepared by Educational Testing Service (ETS) (Anderson, Jenkins, & Miller, 1996).

A number of new complexities were planned for the 1996 assessment, such as increased use of manipulatives in mathematics, theme blocks in mathematics, hands-on tasks in science, and performance items in dance, music, theater, and visual arts. The complexities of mathematics and science substantially increased the scope of the 1996 assessment, as originally defined, and were rehearsed as part of the field test.

In September 1994, letters were sent from the U.S. Department of Education to all Chief State School Officers inviting them to participate in the 1995 field test of materials and

[1] Nancy W. Caldwell directed survey operations and field activities for the NAEP assessments.

procedures. In an effort to secure the participation of more schools and to lessen the burden of participation on jurisdictions, ETS and Westat offered to perform all of the work involved, including sampling, communicating with school staff, and administering the assessment.

The school sample for the field test included both public and nonpublic schools and was designed to involve as many states as possible, thus limiting the burden on each state. However, states with small numbers of schools in which all schools were already involved in the 1994 National Assessment program were excluded from the field test sample. As a result, the original field test sample consisted of 1,129 public and nonpublic schools spread roughly in proportion to the population across 38 states. Because the states' responsibilities were very limited in the field test, they were asked only to notify districts of their inclusion, and to indicate their support for participation in the field test. Schools selected for the 1995 field test were designated to have either arts sessions or mathematics and science sessions, but not both.

Because the focus of the field test was to have as many schools participate as possible, flexibility was allowed in substituting for the original selections. Three forms of substitution were available to replace sampled schools that did not participate in the field test. The first type were schools identified by Westat and located within the same district as the originally sampled schools. These substitute schools were demographically comparable to those in the original sample. A second school substitution option allowed district superintendents to choose their own alternate school. In the event that a district refused to participate, the third option was an "out of district" substitute, identified by Westat. The type and number of sessions scheduled for an originally selected school were carried over to the substitute school.

During the period from October to December 1994, all districts and schools in the field test sample were contacted, cooperation secured, and assessment schedules set. To accomplish these initial tasks, 21 of the most experienced NAEP supervisors were trained during a three-day session (in early October 1994) conducted by Westat project staff. Following training, each of the supervisors was responsible for scheduling activities in several states. In December 1994, the NAEP field staff was expanded to 72 supervisors. All supervisors, including those in the original group, attended the second training session. After opening plenary sessions, the trainees were divided into two groups: arts and mathematics/science. Because of the complicated nature of the arts field test, it was decided to have supervisors specialize in the administration of either arts or mathematics/science sessions. Training focused on a review of the scheduling activities during the fall (e.g., results of initial contacts with districts and schools); sampling procedures; preparation and distribution of school, teacher, and student questionnaires; administration of the performance-based arts tasks; classroom management techniques; exercise administrator training; and completion of administrative forms and procedures.

The period from January 2-20, 1995, was set aside for supervisors to call and visit the schools in their assignments, draw student samples, prepare Administration Schedules, and prepare and distribute teacher, school, and SD/LEP student questionnaires. Assessments were conducted during the period from January 23 through March 10, 1995. Mathematics and science sessions were scheduled to be completed by February 24, and arts sessions continued through the end of the data collection period. Throughout the field testing period, supervisors reported directly to Westat's field director through six field managers.

4.2.2 Results of the Field Test

A total of 963 originally selected schools and alternates actually participated in the field test. The final assessed sample of schools included 434 schools at grade 4, 395 schools at grade 8, and 134 schools at grade 12.

A total of 46,514 students participated in the field test. Of this number, 17,212 students participated in the 1995 arts field test; these students will be discussed in a later report on the arts assessment. Student participation in mathematics and science included 11,014 students at grade 4, 11,641 students at grade 8, and 6,647 students at grade 12.

4.3 THE 1996 STATE ASSESSMENT

Forty-four states, the District of Columbia, and Guam volunteered for the 1996 State Assessment, as did the Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) and the Department of Defense Dependents Schools Office of Dependents Education (DoDDS). Table 4-1 identifies the jurisdictions participating in the State Assessment.

Table 4-1
Jurisdictions Participating in the 1996 State Assessment Program in Science[1]

Jurisdictions			
Alabama	Georgia	Mississippi	Rhode Island
Alaska	Guam	Missouri	South Carolina
Arizona	Hawaii	Montana	Tennessee
Arkansas	Indiana	Nebraska	Texas
California	Iowa	Nevada	Utah
Colorado	Kentucky	New Hampshire	Vermont
Connecticut	Louisiana	New Jersey	Virginia
Delaware	Maine	New Mexico	Washington
DoDEA/DDESS	Maryland	New York	West Virginia
DoDEA/DoDDS	Massachusetts	North Carolina	Wisconsin
District of Columbia	Michigan	North Dakota	Wyoming
Florida	Minnesota	Oregon	

[1]The 1996 State Assessment in science was conducted at grade 8 only, except for Department of Defense Education Activity (DoDEA) schools that were also assessed at grade 4 (see Section 4.3.5) as part of a special assessment.

4.3.1 Overview of Responsibilities

Data collection for the 1996 State Assessment involved a collaborative effort between the participating jurisdictions and the NAEP contractors, especially Westat, the field administration contractor. Westat's responsibilities included:

- selecting the sample of schools and students for each participating jurisdiction;
- developing the administration procedures and manuals;
- training state personnel to conduct the assessments; and
- conducting an extensive quality assurance program.

Each jurisdiction volunteering to participate in the 1996 program was asked to appoint a state coordinator. In general, the coordinator was the liaison between NAEP/Westat staff and the participating schools. In particular, the state coordinator was asked to:

- gain the cooperation of the selected schools;
- assist in the development of the assessment schedule;
- receive the lists of all grade-eligible students from the schools;
- coordinate the flow of information between the schools and NAEP;
- provide space for the Westat state supervisor to use when selecting the sample of students;
- notify assessment administrators about training and send them their manuals; and
- send the lists of sampled students to the schools.

At the school level, an assessment administrator was responsible for preparing for and conducting the assessment session(s) in one or more schools. These individuals were usually school or district staff and were trained by Westat staff. The assessment administrator's responsibilities included:

- receiving the list of sampled students from the state coordinator;
- identifying sampled students who should be excluded;
- distributing assessment questionnaires to appropriate school staff and collecting them upon their completion;
- notifying sampled students and their teachers;
- administering the assessment sessions(s);
- completing assessment forms; and
- preparing the assessment materials for shipment.

Decisions on exclusion were made in consultation with school staff and were guided by the SD/LEP questionnaires completed by the school staff.

Westat also hired and trained a state supervisor for each jurisdiction. The 1996 State Assessment involved about the same number of state supervisors (Westat staff) as both the 1992 and 1994 assessments, since approximately the same number of jurisdictions were involved each year. In addition, three troubleshooters were trained in case any state supervisor was unable to complete their assignment. The primary tasks of the state supervisor were to:

- select the samples of students to be assessed;
- recruit and hire the quality control monitors throughout their jurisdiction;
- conduct in-person assessment administration training sessions; and
- coordinate the monitoring of the assessment sessions and makeup sessions.

Westat hired and trained six field managers for the State Assessment. Each field manager was responsible for working with the state coordinators of seven to eight jurisdictions and for overseeing assessment activities. The primary tasks of the field managers were to:

- obtain information about cooperation and scheduling;
- make sure the arrangements for the assessments were set and assessment administrators identified; and
- schedule the assessment administrators training sessions.

In addition, Westat hired between four and six quality control monitors in each jurisdiction to monitor assessment sessions.

4.3.2 Schedule of Data Collection Activities

Mid-September 1995	Westat sent lists of sampled schools for the national and state assessments and informational materials to the state coordinators.
October 1995	Westat field managers visited individual jurisdictions to explain the computerized state coordinator system, which was used to keep track of assessment-related activities.
September - December 1995	State coordinators obtained cooperation from districts and schools. State coordinators reported participation status to Westat field managers via computer files or printed lists. State coordinators sent student listing forms and supplemental student listing forms to participating schools.
November 9 - 12, 1995	State supervisor training.
November 17, 1995	Suggested cutoff for decisions on school participation and submission of lists of grade-eligible students to state coordinators for sampling purposes.

December 4 - 15, 1995	NAEP supervisor visited state coordinators' offices to select student samples and prepare Administration Schedules listing the students selected for each sample.
December 15, 1995	Westat delivered training session schedule and copies of assessment administrator manuals to state coordinators for distribution.
December 15, 1995 - January 15, 1996	State coordinators notified assessment administrators of the date, time, and location of training and sent each a copy of the manual for assessment administrators.
January 4 - 6, 1996	Training session for quality control monitors.
January 9 - 26, 1996	Supervisors conducted assessment administrator training sessions throughout respective jurisdictions.
January 29 - March 1, 1996	Assessments conducted and monitored.
March 4 - 8, 1996	Makeup week for rescheduled assessments or completed assessments requiring makeup.

4.3.3 Preparations for the State Assessment

The focal point of the schedule for the State Assessment was the period between January 29 and March 4, 1996, when the assessments were conducted in the schools. However, as with any undertaking of this magnitude, the project required many months of planning and preparation.

Westat selected the samples of schools according to the procedures described in Chapter 3. In mid-September 1995, lists of the selected schools and other materials describing the State Assessment program were sent to state coordinators. Most state coordinators preferred that NAEP provide a suggested assessment date for each school. School listings were updated with this information and were sent to the state coordinators, along with other descriptive materials and forms, by December.

State coordinators were also given the option of receiving the school information in the form of a computer database with accompanying management information software. This system enabled state coordinators to keep track of the cooperating schools, the assessment schedule, the training schedule, and the assessment administrators. Coordinators could choose to receive a laptop computer and printer or to have the system installed on their own computer. Westat field managers traveled to the state offices to explain the computer system to the state coordinators and their staff. Only one jurisdiction chose not to use the computerized system. In this case, the state coordinator kept track of information on logs and lists provided by Westat. This printed information was mailed to the field manager and dictated during a regularly scheduled telephone conversation. The field manager then entered the data into the computer database, the data were transmitted to Westat, and reports were produced.

Six of the most experienced NAEP supervisors served as field managers, the primary link between NAEP and the state coordinators. During late summer and early fall 1995, the field managers received copies of all materials sent to state coordinators, developed a preliminary Assessment Schedule for all schools in their jurisdictions, and became thoroughly familiar with the computer system. As liaisons with the state coordinators, they visited each jurisdiction to train staff in the use of the computer system. Later in the project schedule, they attended training sessions for the supervisors and quality control monitors and also presented some of the training material at each of these sessions.

The field managers used the same computer system as the state coordinators to keep track of the schools and the schedule. The state coordinators sent updates via computer disks, telephone, or print to their field manager, who then entered the information into the system. Weekly transmissions were made from the field manager to Westat.

By November, Westat had hired one state supervisor for each participating jurisdiction. The state supervisors attended a training session held November 9-12, 1995. This training session focused on the state supervisors' immediate tasks — selecting the student samples and hiring quality control monitors. Supervisors were given the training script and materials for the assessment administrators' training sessions they would conduct in January so they could become familiar with these materials.

The state supervisors' first task after training was to complete the selection of the sample of students who were to be assessed in each school. All participating schools were asked to send a list of their grade-eligible students to the state coordinator by November 17. Sample selection activities were conducted in the state coordinator's office unless the state coordinator preferred that the lists be taken to another location.

Using a preprogrammed calculator, the supervisors generally selected a sample of 30 students per session type per school with three exceptions: in schools with fewer than 30 students in the grade to be assessed, all of the students were selected; in schools in which more than one session was scheduled, 60 students (or some multiple of 30 students) were selected; and in schools with no more than 33 students in the grade, all students were selected for the assessment.

After the sample was selected, the supervisor completed an Administration Schedule for each session, listing the students to be assessed. The Administration Schedules for each school were put into an envelope and given to the state coordinator to send to the school two weeks before the scheduled assessment date. Included in the envelope were instructions for sampling students who had enrolled at the schools since the creation of the original list.

During the months of November and December 1995, the state supervisors also recruited and hired quality control monitors to work in their jurisdictions. It was the quality control monitor's job to observe the sessions designated to be monitored, to complete an observation form on each session, and to intervene when the correct procedures were not followed. Because earlier results indicated little difference in performance between monitored and unmonitored schools, and in an effort to reduce costs, the percentage of public schools to be monitored was maintained at 25 percent (i.e., the reduced monitoring rate initiated in 1994). The monitoring rate for nonpublic schools was also reduced to 25 percent (from 50% in 1994, which was the first year that nonpublic schools were assessed by NAEP). As has been customary in the past, monitoring was conducted at 50 percent for jurisdictions that were new to the State Assessment in 1996. The schools to be monitored were known only to contractor staff; it was not indicated on any of the listings provided to state staff.

Approximately 400 quality control monitors were trained in a session held in early January 1996. The first day of the training session was devoted to a presentation of the assessment administrators' training program by the state supervisors, which not only gave the monitors an understanding of what assessment administrators were expected to do, but gave state supervisors an opportunity to practice presenting the training program. The remaining days of the training session were spent reviewing the quality control monitor observation form and the role and responsibilities of the quality control monitors.

Almost immediately following the quality control monitor training, supervisors began conducting training for assessment administrators. Each quality control monitor attended at least two training sessions, to assist the state supervisor and to become thoroughly familiar with the assessment administrator's responsibilities. Most jurisdictions had approximately 14 training sessions in which approximately 217 assessment administrators were trained. Almost 10,400 assessment administrators were trained by the time assessments began on January 29, 1996.

To ensure uniformity in the training sessions, Westat developed a highly structured program involving a script for trainers, a videotape, and an example to be completed by the trainees. The training package, developed for previous state assessments, was revised to reflect the subjects and grades assessed in 1996. The supervisors were instructed to read the script verbatim as they proceeded through the training, ensuring that each trainee received the same information. The script was supplemented by the use of overhead transparencies, displaying the various forms that were to be used and enabling the trainer to demonstrate how they were to be filled out.

The videotape was also revised from previous versions to include information about assessing both fourth- and eighth-grade students. The 1996 version of the video ran just over one hour.

All of the information presented in the training session was included in Westat's *Manual for Assessment Administrators*. Copies of the manual were sent by Westat to the state coordinators by December 15, 1995, so that they could be distributed to the assessment administrators before the training sessions. The method of distribution and the amount of time that the assessment administrators had to study the manual probably varied from jurisdiction to jurisdiction. The majority of the assessment administrators appeared to have become at least somewhat familiar with the manual prior to their training. The training stressed that answers to all questions about procedures or forms could be found in the manual. In addition, assessment administrators were provided with a toll-free number that could be used to contact Westat if they had any procedural questions or were in need of additional materials. During the assessment period, this telephone number was used extensively.

The entire training session generally ran for about one-half day until 2 p.m. including lunch.

4.3.4 Monitoring of Assessment Activities

Two weeks prior to the scheduled assessment date, the assessment administrator received the Administration Schedule and assessment questionnaires and materials. Five days before the assessment, the quality control monitor made a call to the administrator and recorded the results of the call on the Quality Control Form for Monitored Schools because the assessment administrators were not supposed to know in advance which sessions were designated to be monitored. The pre-assessment call was conducted in exactly the same way regardless of whether the school was to be monitored or not. For example, directions to the school were obtained even if the school was in the unmonitored sample. Most of the questions asked in the pre-assessment call were designed to gauge whether the assessment administrator had received all materials needed and had completed the preparations for the assessment.

If the sessions in a school were designated to be monitored, the quality control monitor was to arrive at the school one hour before the scheduled beginning of the assessment to observe preparations for the assessment. To ensure the confidentiality of the assessment items, the booklets were packaged in shrink-wrapped bundles and were not to be opened until the quality control monitor arrived or 45 minutes before the session began, whichever occurred first.

In addition to observing the opening of the bundles, the quality control monitor used the Quality Control Form to check that the following had been done correctly: sampling newly enrolled students, reading the script, distributing and collecting assessment materials, timing the booklet sections, answering questions from students, and preparing assessment materials for shipment. After the assessment was over, the quality control monitor obtained the assessment administrator's opinions of how the session went and how well the materials and forms worked. The 14-section booklet, Quality Control Form for Monitored Schools, is included in the *Report on Data Collection Activities for the 1996 National Assessment of Educational Progress* (Westat, Inc., 1996).

If four or more students were absent from the session, a makeup session was to be held. If the original session had been monitored, the makeup session was also monitored. This required coordination of scheduling between the quality control monitor and assessment administrator.

4.3.5 Participation of Department of Defense Education Activity Schools (DoDEA)

The schools run by the Department of Defense at military bases and other installations around the world participated in the NAEP State Assessment for the second time in 1996. The participation of the selected schools was mandated by DoDEA. To accommodate the geographic diversity of DoDEA schools, some minor adaptations were made in the preparatory activities used for the other jurisdictions.

The data collection in DoDEA schools was expanded in 1996 so that both DDESS and DoDDS schools were surveyed. In 1994, only the schools at overseas installations were sampled as part of the State Assessment. Also, DoDEA chose to conduct science assessments at grade 4 (in other State NAEP schools, science was conducted only at grade 8) so that both mathematics and science data were collected at both grades 4 and 8 in DoDEA schools.

Many of the quality control monitors hired for the DoDEA schools were based overseas, and many had previous experience working within the DoDEA system. They were referred to Westat by DoDEA. All quality control monitors for the DoDEA schools attended the quality control training in Los Angeles and several assessment administrator training sessions in the geographic areas in which they worked.

The samples of students to be assessed in the DoDEA schools were selected in the Westat home office, using standard NAEP procedures, from lists of students produced in the DoDEA offices in Northern Virginia. Due to privacy concerns, only student ID numbers and not student names appeared on the DoDEA lists. Thus, after sampling, the Administration Schedules contained only the ID numbers, and the assessment administrators consulted school records and added the names of the students to the Administration Schedules prior to the assessments.

Two field supervisors were hired specifically to conduct assessment administrator trainings and monitor quality control monitors in the DoDDS schools. The DoDEA liaison in Northern Virginia, who essentially functioned as the state coordinator, arranged the assessment administrator training sessions, all of which were held in schools or other facilities on the bases. In many cases, the quality control monitors were required to obtain special clearances through DoDEA to visit the bases for training and the assessments.

The assessments in DoDEA schools were conducted using the same procedures as in all State Assessment schools with the one exception that DoDEA included science assessments at both grades 4 and 8.

4.3.6 Exclusion of Students from the Assessment

Due to recent interest in including as many students as possible in NAEP and other educational assessments, efforts were initiated in the 1995 field test to explore the impact of redefining the NAEP inclusion criteria for students with disabilities and/or limited English proficiency (SD/LEP). This investigation was continued in 1996 in both the national and State Assessments.

The approach taken in the 1996 State Assessment was to divide the school sample into two, equal-size subsamples, referred to as S1 and S2. The schools in the S1 subsample were asked to apply the “old” (used in previous years) inclusion criteria; the S2 schools received a “new,” revised criteria. The assessment administration for a school assured that the appropriate set of inclusion criteria were used in each school. Training of each member of the field staff included information about the two sets of inclusion criteria. Figures 4-1 and 4-2 describe the criteria for S1 and S2.

Figure 4-1
S1 Criteria

A student identified on the Administration Schedule as LEP may be excluded from the assessment if he or she:

1. is a native speaker of a language other than English,
2. has been enrolled in an English-speaking school (not including a bilingual education program) for less than two years, and
3. is judged to be incapable of taking part in the assessment.

A student identified on the Administration Schedule as SD or an equivalent classification may be excluded from the assessment if:

1. the student is mainstreamed less than 50 percent of the time in academic subjects and is judged incapable of participating meaningfully in the assessment, or
2. the Individualized Education Plan (IEP) team or equivalent group has determined that the student is incapable of participating meaningfully in the assessment.

SD/LEP students meeting the above criteria should be assessed if, in the judgment of school staff, they are capable of taking the assessment.

Figure 4-2
S2 Criteria

A student who is identified on the Administration Schedule as LEP and who is a native speaker of a language other than English should be included in the NAEP assessment unless:

1. the student has received mathematics, science, and language arts instruction primarily in English for less than three school years, including the current year, or
2. the student cannot demonstrate his or her knowledge of mathematics or science in English without an accommodation or adaptation.

A student identified on the Administration Schedule as SD or an equivalent classification should be included in the NAEP assessment unless:

1. the IEP team or equivalent group has determined that the student cannot participate in assessments such as NAEP,
2. the student's cognitive functioning is so severely impaired that she or he cannot participate, or
3. the student's IEP requires that the student be tested with an accommodation or adaptation and the student cannot demonstrate his/her knowledge of mathematics or science without that accommodation or adaptation.

The school person most knowledgeable about each student classified as IEP or LEP should complete an SD/LEP Questionnaire about the student.

The preliminary, unweighted proportion of students in the S1/S2 subsampling suggest that applying the new (S2) or old (S1) criteria result in virtually no change in the proportions of students excluded from the NAEP assessments as SD or LEP. For example, in grade 4 public schools, in both subsamples, about 5.6 - 5.7 percent of the students were excluded as SD and about 1.6 - 1.7 percent were consistently excluded as LEP students. The rates are slightly lower for grade 8 public-school students—just below five percent for SD exclusions, and about one percent for LEP — and again consistent across the two subsamples. The rates for nonpublic schools were lower still, that is, consistently less than half the size of the public-school rates and very similar across the S1 and S2 subsamples.

4.3.7 School and Student Participation

Table 4-2 shows the results of the state coordinators' efforts to gain the cooperation of the selected schools. Overall, for the 1996 State Assessment in science, 3,926 public schools and 474 nonpublic schools for grade 8 participated.

Participation results for students in the 1996 State Assessment in science are given in Table 4-3. Over 136,800 eighth-grade students were sampled. As can be seen from the table, the original sample, which was selected by the NAEP state supervisors, comprised approximately 134,000 (or 98%) of the total number of students sampled for grade 8. The original sample size was increased somewhat after the supplemental samples had been drawn (from students newly enrolled since the creation of the original lists).

Table 4-2
School Participation, 1996 State Assessment in Science

	Grade 8	
	Public	**Nonpublic**
Number of schools in original sample	4,478	1,063
Number of schools not eligible	113	204
Number of eligible schools in original sample	4,365	859
Non-cooperating (e.g., school, district, or state refusal)	560	277
Cooperating	3,805	582
Number of substitutes provided for non-cooperating schools	408	152
Number of participating substitutes for non-cooperating schools	121	16
Total number of schools participating (after substitution)	3,926	474

Table 4-3
Student Participation, 1996 State Assessment in Science

	Grade 8	
	Public	**Nonpublic**
Number sampled	128,534	8,360
Original sample	125,750	8,277
Supplemental sample	2,784	83
Percent increase in original sample by adding supplemental sample	2.2%	1.0%
Number (%) of originally sampled students withdrawn	5,824 (4.6%)	135 (1.6%)
Number of students excluded[1]	7,452	44
Number (%) of sampled students identified as SD	13,482 (10.5)	179 (2.1%)
Number (%) of sampled students excluded as SD	6,127 (4.8%)	34 (0.4%)
Number (%) of sampled students identified as LEP	2,991 (2.3%)	40 (0.5%
Number (%) of sampled students excluded as LEP	1.464 (1.1%)	10 (0.1%)
Number of students to be assessed	115,258	8,181
Number of students assessed	104,998	7,805
Original sessions	103,549	7,763
Makeup sessions	1,449	42
Student participation rates		
Before makeups	89.8%	94.9%
After makeups	91.1%	95.4%

[1] *To be excluded, a student had to be designated as SD or LEP and judged incapable of participating in the assessment. A student could be identified as both SD and LEP, resulting in this number being less than the sum of the students excluded as SD or LEP.*

4.3.8 Results of the Observations

During the assessment sessions, the quality control monitors observed whether the assessment environment was adequate or inadequate based on factors such as room size, seating arrangements, noise from hallways or adjacent rooms, and lighting. (If the room was unsuitable, however, the quality control monitors did not routinely ask the assessment administrator to make other arrangements.) Of the 3,776 monitored assessment sessions where quality control monitors recorded an observation, the quality control monitors felt that 96 percent of the sessions were held in suitable surroundings.

The *Manual for Assessment Administrators* encouraged assessment administrators to use an assistant during the assessment session, a suggestion that came from the earliest state assessment in 1990. To measure how frequently that advice was heeded, quality control monitors noted whether an assistant was used in the monitored sessions. The results indicate that assistants were used for 60 to 70 percent of the public-school sessions, with the largest percentage (66% - 70%) noted for grade 8 sessions. In nonpublic schools, however, an assistant was employed less often (29% - 40% of the time), which is possibly a reflection of fewer staff resources and generally smaller session sizes in nonpublic schools. Assessment administrators used assistants in varying capacities. The *Manual for Assessment Administrators* was very emphatic that only a NAEP-trained person could actually administer the assessment session. Almost always, assistants helped to supervise the session and to prepare, distribute, and collect assessment materials and/or booklets.

The assessment administrators were asked to estimate the total time that they spent on the preparations for and the conduct of the assessment, including their attendance at the training session. Estimates for 1996 were similar to those for 1992 because two subjects were assessed in each of these years (compared to 1994 when only one subject was assessed). In 1996, a majority of the assessment administrators with grade 4 sessions (63% in public and 82% in nonpublic schools) stated that they spent less than 20 hours on the assessment. For grade 8, however, only 30 percent of the assessment administrators in public schools, compared to 73 percent of those in nonpublic schools, spent fewer than 20 hours. The variation in time distribution for grade 8 public schools, particularly compared to public schools at grade 4, is most likely due to the fact that two session types (mathematics and science) were usually conducted by each grade 8 assessment administrator, but only one session type (mathematics) was held at grade 4. This does not appear to hold true for nonpublic schools, however, where the distribution of time spent is more similar for grades 4 and 8. It is evident that assessment administrators in nonpublic schools spent fewer hours overall on the assessment than did assessment administrators in public schools. Potential explanations might be the generally smaller sessions sizes in nonpublic schools (i.e., fewer materials to prepare and ship) and the possibility that some grade 8 schools may have used more than one assessment administrator with each assessment administrator conducting one session (but compiling a larger total time for all sessions combined).

Quality control monitors reported that they observed the opening of assessment booklet bundles for 3,539 (or 89%) of the monitored sessions, and it is assumed that these bundles were opened at the proper time. In two percent of the sessions, however, the bundle opening was not observed due to quality control monitor error, (e.g., the quality control monitor was late, in the wrong place, or miscommunicated with the assessment administrator); presumably, some (or probably most) of these bundles were opened at the correct time. For another two percent of sessions, the quality control monitors were unable to observe the bundle opening that occurred early due to assessment administrator error (e.g., the assessment administrator misunderstood the procedures, felt more time was needed, had scheduling conflicts, or needed to prepare for multiple sessions starting at the same time). Information on the opening of the assessment booklet bundles was not reported for the remaining seven percent of the monitored sessions.

When queried, the quality control monitors felt most positive about the attitudes of the assessment administrators and somewhat less positive about the attitudes of other school staff and the students towards the assessment.

Quality control monitors concluded the summary section by assigning a final rating of the assessment administrator's performance. With this rating, the quality control monitor reconsidered the session from the vantage point of how well it would have gone without the quality control monitor's presence. Eighty-four percent of the assessment administrators in monitored sessions were self-reliant or needed to consult the quality control monitors for only one or two minor items. Only about four or five percent had serious difficulty conducting the session (that is, relied on the quality control monitor to initiate procedures or conduct the session).

After the conclusion of the assessment sessions, Westat mailed state coordinators a short survey to obtain their reactions to the operations associated with the 1996 State Assessment and any suggestions they had for improving the program. Thirty-seven state coordinators responded by returning the survey or by providing their responses over the telephone. A detailed summary of the state coordinators' responses is contained in the *Report on Data Collection Activities for All States* (Westat, Inc., 1996), which was distributed to state coordinators in October 1996. Some of the responses from the state coordinators included:

- Fifteen of the 37 reporting jurisdictions mandated participation in the 1996 State Assessment;

- Seven jurisdictions reported that they helped gain the cooperation of nonpublic schools. Most of these provided a letter from the state superintendent of schools, and others answered questions.

- Twenty-nine jurisdictions used the computer system throughout the field testing period. Seven jurisdictions used the system initially but not necessarily during the assessment period, and one jurisdiction did not use the system at all. The jurisdictions seemed to be comfortable with the computer system and were able to use it effectively. A fairly common suggestion was to expand the documentation and capabilities regarding label production.

- Of the jurisdictions reporting on staff time devoted to NAEP, state coordinators spent an average of 34 days (ranging from 2 to 100 days) on NAEP activities, and other staff spent an average of 28 days (ranging from 2 to 85 days).

- Reactions to the 1996 State Assessment were quite positive. Twenty-five of the 28 state coordinators who expressed an opinion said that the assessments went "very well" or "well" — even though this was a challenging year in terms of bad weather, missed instruction time, and school staff burden.

Chapter 5

PROCESSING AND SCORING ASSESSMENT MATERIALS[1]

Patrick B. Bourgeacq, Charles L. Brungardt, and Brent Studer
National Computer Systems

5.1 OVERVIEW

This chapter reviews the processing and scoring activities conducted by National Computer Systems (NCS) for the 1996 NAEP State Assessment. The 1996 assessment presented the greatest challenge in processing and scoring NAEP data to date. For this assessment, NCS was charged with processing and scoring the largest assessment in the history of NAEP in the shortest amount of time. Further, image scanning processes, eliminating almost all paper handling during scoring and improving monitoring and reliability scoring, increased to nearly twice that of the 1994 assessment. In the early 1990s, NCS developed and implemented flexible, innovatively designed processing programs and a sophisticated Process Control System that allows the integration of data entry and workflow management systems to accomplish this work.

This chapter begins with a description of the various tasks performed by NCS, detailing printing, distribution, receipt control, scoring, and processing activities. It also discusses specific activities involved in processing the assessment materials, and presents an analysis of several of those activities. The chapter provides documentation for the professional scoring effort — scoring guides, training papers, papers illustrating sample score points, calibration papers, calibration bridges, and interreader reliability reports. The detailed processing specifications and documentation of the NAEP Process Control System are presented in the final sections of the chapter.

5.1.1 Innovations for 1996

Much of the information necessary for documentation of accurate sampling and for calculating sampling weights is collected on the Administration Schedules that, until 1993, were painstakingly filled out by hand by Westat administrative personnel. In 1994, for the first time, much of the work was computerized — booklets were preassigned and booklet ID numbers were preprinted on the Administration Schedule. When Westat personnel received the documents, they filled in only the “exception” information. This new method also permitted computerized updating of information when the Administration Schedules were received at NCS, eliminating the need to sort and track thousands of pieces of paper through the processing stream.

The introduction of image processing and image scoring further enhanced the work of NAEP. Image processing and scoring were successfully piloted in a side-by-side study conducted during the 1993 NAEP field test, and so became the primary processing and scoring methods for the 1994 and 1996 State Assessments. Image processing allowed the automatic collection of handwritten demographic data from the administrative schedules and the student test booklet

[1] Patrick Bourgeacq is the project director for scoring. All of the authors were involved in the processing and scoring procedures for the NAEP State Assessments.

covers through intelligent character recognition (ICR). This service was a benefit to the jurisdictions participating in NAEP because they were able to write rather than grid certain information — a reduction of burden on the schools. Image processing also made image scoring possible, eliminating much of the time spent moving paper as part of the scoring process. The images of student responses to be scored were transmitted electronically to the scoring center, located at a separate facility from where the materials were processed. This process enhanced the reliability and monitoring of scoring and allowed both NCS and ETS to focus attention on the intellectual process of scoring student responses (Johnson, 1994).

Tables 5-1 and 5-2 give an overview of the processing volume and the schedule for the 1996 NAEP State Assessment in science.

Table 5-1
1996 NAEP State Assessment in Science Processing Totals

Document/Category	Totals
Number of sessions	15,487
Assessed student booklets	356,447
Absent student booklets	27,743
Excluded student booklets	25,713
SD/LEP questionnaires	47,708
School questionnaires	9,470
Teacher questionnaires	39,311
Scanned documents	356,447
Scanned sheets	9,829,970
Key-entered documents[1]	0

[1]No Braille booklets and only one large-print booklet were received from the 1995 field test. Rather than key enter only one booklet, the decision was made to bypass the key-entry stage and let the scoring center score it directly from the booklet. Thus, there were zero key-entered documents.

Table 5-2
1996 NAEP State Assessment, NCS Schedule

Activity	Planned Start Date	Planned Finish Date	Actual Start Date	Actual Finish Date
Printing	9/2/95	12/11/95	9/2/95	12/11/95
Grade 8 Teacher Questionnaires delivered to NCS	10/12/95	10/12/95	10/16/95	10/16/95
Administration Schedule delivered to NCS	10/18/95	10/18/95	10/23/95	10/23/95
Grade 8 School Characteristics and Policies Questionnaires at NCS	10/20/95	10/20/95	10/23/95	10/23/95
SD/LEP Roster delivered to NCS	10/20/95	10/20/95	10/24/95	10/24/95
Grade 4 School Characteristics and Policies Questionnaires at NCS	10/20/95	10/20/95	10/25/95	10/25/95
Pre-packaging begins	10/23/95	12/20/95	10/16/95	12/1/95
Grade 8 science Teacher Questionnaires at NCS	10/30/95	10/30/95	11/1/95	11/1/95
NCS/ETS meet to review items and scoring schedule	11/2/95	11/3/95	11/2/95	11/3/95
Grade 8 science spiral material at NCS	11/6/95	11/13/95	11/13/95	11/21/95
State supervisor training	11/9/95	11/12/95	11/9/95	11/11/95
Administration Schedule address file from Westat	11/20/95	11/20/95	11/22/95	11/22/95
95% session data file of schools from Westat	11/22/95	11/22/95	11/22/95	1/5/96
SD/LEP Questionnaire delivered to NCS	11/22/95	11/22/95	12/5/95	12/11/95
Grade 4 science spiral material at NCS	11/22/95	11/30/95	11/21/95	12/1/95
Print Administration Schedules	11/27/95	11/27/95	11/28/95	10/23/95
Ship Administration Schedules to Westat state supervisors	11/29/95	11/29/95	11/28/95	1/26/96
All materials at NCS for packaging	11/29/95	12/1/95	12/1/95	12/15/95
State supervisor training materials shipped	12/15/95	12/15/95	12/13/95	12/13/95
School address file from Westat	12/18/95	12/18/95	11/29/95	11/29/95
Final packaging	12/26/95	2/3/96	12/26/95	2/7/96
Receiving	1/30/96	3/5/96	2/6/96	3/12/96
Processing	2/2/96	3/22/96	2/6/96	4/5/96
PSC selects science table leaders	3/1/96	3/1/96	2/1/96	2/28/96
Scoring training preparation	3/4/96	3/22/96	3/4/96	3/22/96
Scorers assigned to teams	3/11/96	3/11/96	3/11/96	3/11/96
Training and scoring	3/18/96	5/31/96	3/18/96	6/7/96

Table 5-2 (continued)
1996 NAEP State Assessment, NCS Schedule

Activity	Planned Start Date	Planned Finish Date	Actual Start Date	Actual Finish Date
Weights data shipped - science grade 8	3/30/96	4/1/96	4/8/96	4/24/96
Weights data shipped - grade 4	3/15/96	3/18/96	3/28/96	4/22/96
Grade 4 science data tape sent to ETS	5/31/96	5/31/96	5/30/96	5/30/96
Grade 8 science data tape sent to ETS	5/31/96	5/31/96	6/26/96	6/26/96
School Characteristics and Policies Questionnaires data tape shipped to ETS	7/11/96	7/12/96	7/11/96	7/11/96
Teacher Questionnaires data tape shipped	7/18/96	7/19/96	7/24/96	7/24/96
SD/LEP Questionnaires data shipped to ETS	7/26/96	7/29/96	8/7/96	8/7/96

5.2 PRINTING

5.2.1 Overview

For the 1996 NAEP assessments, 255 unique documents were designed. NCS printed more than 1,900,000 booklets and forms, totaling over 58 million pages.

Printing preparations began with the design of the booklet covers in June 1995. This was a collaborative effort involving staff from ETS, Westat, and NCS. Because the goal was to design one format for use with all of the booklets, necessary data elements to be collected for the different assessment types had to be agreed upon. In a similar collaboration with ETS and Westat, NCS prepared Administration Schedules and questionnaire rosters, and the camera-ready copies for the documents were created and edited. The printing of assessment booklets, questionnaires, and tracking forms for the main and state assessments was complete by December 11, 1995.

5.2.2 State Assessment Printing

Camera-ready data for all of the science blocks were created by ETS, as were some of the directions and all of the background blocks. Because large numbers of documents were to be printed in a relatively short period of time, preliminary composition work was begun by the NCS printer in Columbia, Pennsylvania, and the required numbers of negatives for each booklet component were made. Performing these preliminary tasks was crucial to meeting the delivery schedule.

The actual assembly of booklets began after all parts needed for a particular booklet were received and the Office of Management and Budget (OMB) had given its approval to print. ETS supplied booklet maps that specified the order of blocks in each booklet (see Chapter 2, Table 2-5, for the contents of each booklet). Using these booklet maps and mock-ups of booklets as guides, the NCS printer assembled prepared negatives into complete booklets. Generally, five weeks elapsed between receipt of final copy and delivery of printed booklets.

The printer forwarded proofs of the booklets and questionnaires to ETS and to NCS for review and approval to print. Clean-up work and changes, where necessary, were indicated on the proofs, which were returned to the printer. Once approved, the booklets were printed.

As the booklets and forms were printed, pallets of documents were received and entered into NCS's Inventory Control system. Sample booklets were selected and quality-checked for printing and collating errors. All printing for the 1996 NAEP State Assessment in science was completed by December 11, 1995.

5.3 PACKAGING AND SHIPPING

5.3.1 Distribution

The distribution effort for the 1996 NAEP State Assessment involved packaging and mailing documents and associated forms and materials to individual schools. The NAEP Materials Distribution System, initially developed by NCS in 1990 to control shipments to the schools and supervisors, was utilized again in 1996. Files in the system contained the names and addresses for shipment of materials, scheduled assessment dates, and a listing of all materials available for use by a participant. Changes to any of this information were made directly in the distribution file either manually or via file updates provided by Westat. Figure 5-1 illustrates the process flow for the accountability system and online bundle assignment and distribution system utilized for NAEP.

Bar code technology continued to be utilized in document control. To identify each document, NCS utilized a unique ten-digit numbering system. This numbering system consisted of the three-digit booklet number or form type, a six-digit sequential number, and a check digit. Each form was assigned a range of ID numbers. Bar codes reflecting this ID number were applied to the front cover of each document by NCS bar code processes and high-speed ink jet printers.

Once all booklets from a subject area were bar coded, they were spiraled and bundled into groups of eleven documents. For State Assessment samples in science, NCS spiraled the booklets according to the pattern dictated by ETS in the bundle maps. Booklets were spiraled in such a manner that each booklet appeared in the first position in a bundle approximately the same number of times and that the booklets were evenly distributed across the bundles. This assured that sample sizes of individual booklet types would not be jeopardized if entire bundles were not used.

All booklets had to be arranged in the exact order listed on the bundle header sheet. To ensure the accuracy of each bundle and the security of the NAEP assessment, a quality control plan was utilized to verify the document order of each bundle and to account for all booklets. All bundles that contained a bundle slip were taken to a bar code reader/document transport machine where they were scanned to interpret each bundle's bar code. The file of scanned bar codes was then transferred from the personal computer connected to the scanner to a mainframe data set.

Table 5-3
Documents Printed for the 1996 NAEP State Assessment in Science

Sample	Grade	Document (Booklet)[1]	Subject	Type	Number of Pages	Final Copy from ETS	Approval to Print	Printed Documents Received	Quantity Printed
Main/State	8	S201E	Science	Image scan	48	9/25/95	10/13/95	11/13/95	7,349
Main/State	8	S202F	Science	Image scan	48	9/25/95	10/13/95	11/14/95	7,046
Main/State	8	S203G	Science	Image scan	44	9/25/95	10/16/95	11/14/95	7,352
Main/State	8	S204D	Science	Image scan	44	9/25/95	10/16/95	11/14/95	7,220
Main/State	8	S205E	Science	Image scan	48	9/25/95	10/23/95	11/21/95	8,100
Main/State	8	S206F	Science	Image scan	48	9/25/95	10/16/95	11/14/95	7,140
Main/State	8	S207G	Science	Image scan	44	9/25/95	10/16/95	11/13/95	7,100
Main/State	8	S208D	Science	Image scan	44	9/25/95	10/16/95	11/14/95	7,020
Main/State	8	S209E	Science	Image scan	48	9/25/95	10/16/95	11/13/95	7,120
Main/State	8	S210F	Science	Image scan	48	9/25/95	10/16/95	11/14/95	7,180
Main/State	8	S211G	Science	Image scan	44	9/25/95	10/23/95	11/16/95	7,294
Main/State	8	S212D	Science	Image scan	48	9/25/95	10/23/95	11/14/95	7,340
Main/State	8	S213E	Science	Image scan	44	9/25/95	10/23/95	11/14/95	7,340
Main/State	8	S214F	Science	Image scan	52	9/25/95	10/23/95	11/14/95	7,020
Main/State	8	S215G	Science	Image scan	48	9/25/95	10/23/95	11/14/95	7,200
Main/State	8	S216D	Science	Image scan	52	9/25/95	10/30/95	11/18/95	7,320
Main/State	8	S217E	Science	Image scan	44	9/25/95	10/30/95	11/20/95	7,350
Main/State	8	S218F	Science	Image scan	52	9/25/95	10/30/95	11/20/95	7,300
Main/State	8	S219G	Science	Image scan	44	9/25/95	10/30/95	11/20/95	7,340
Main/State	8	S220D	Science	Image scan	48	9/25/95	10/27/95	11/20/95	9,365
Main/State	8	S221E	Science	Image scan	52	9/25/95	10/23/95	11/17/95	7,340
Main/State	8	S222F	Science	Image scan	52	9/25/95	10/23/95	11/17/95	7,240
Main/State	8	S223G	Science	Image scan	48	9/25/95	10/23/95	11/16/95	7,080
Main/State	8	S224D	Science	Image scan	44	9/25/95	10/23/95	11/14/95	7,352
Main/State	8	S225E	Science	Image scan	48	9/25/95	10/23/95	11/16/95	7,010
Main/State	8	S226F	Science	Image scan	48	9/25/95	10/23/95	11/14/95	7,352

[1] The letters *D, E, F, and G* refer to the ancillary materials that accompanied the assessment booklet.

Table 5-3 (continued)
Documents Printed for the 1996 NAEP State Assessment in Science

Sample	Grade	Document (Booklet)[1]	Subject	Type	Number of Pages	Final Copy from ETS	Approval to Print	Printed Documents Received	Quantity Printed
Main/State	8	S227G	Science	Image scan	44	9/25/95	10/23/95	11/16/95	7,350
Main/State	8	S228D	Science	Image scan	48	9/25/95	10/23/95	11/17/95	7,240
Main/State	8	S229E	Science	Image scan	44	9/25/95	10/23/95	11/17/95	7,352
Main/State	8	S230F	Science	Image scan	48	9/25/95	10/25/95	11/18/95	7,220
Main/State	8	S231G	Science	Image scan	44	9/25/95	10/23/95	11/17/95	7,200
Main/State	8	S232D	Science	Image scan	52	9/25/95	10/23/95	11/17/95	7,280
Main/State	8	S233E	Science	Image scan	44	9/25/95	10/27/95	11/17/95	7,352
Main/State	8	S234F	Science	Image scan	48	9/25/95	10/27/95	11/18/95	7,000
Main/State	8	S235G	Science	Image scan	44	9/25/95	10/30/95	11/18/95	7,240
Main/State	8	S236D	Science	Image scan	48	9/25/95	10/30/95	11/18/95	7,300
Main/State	8	S237E	Science	Image scan	48	9/25/95	10/27/95	11/18/95	7,280

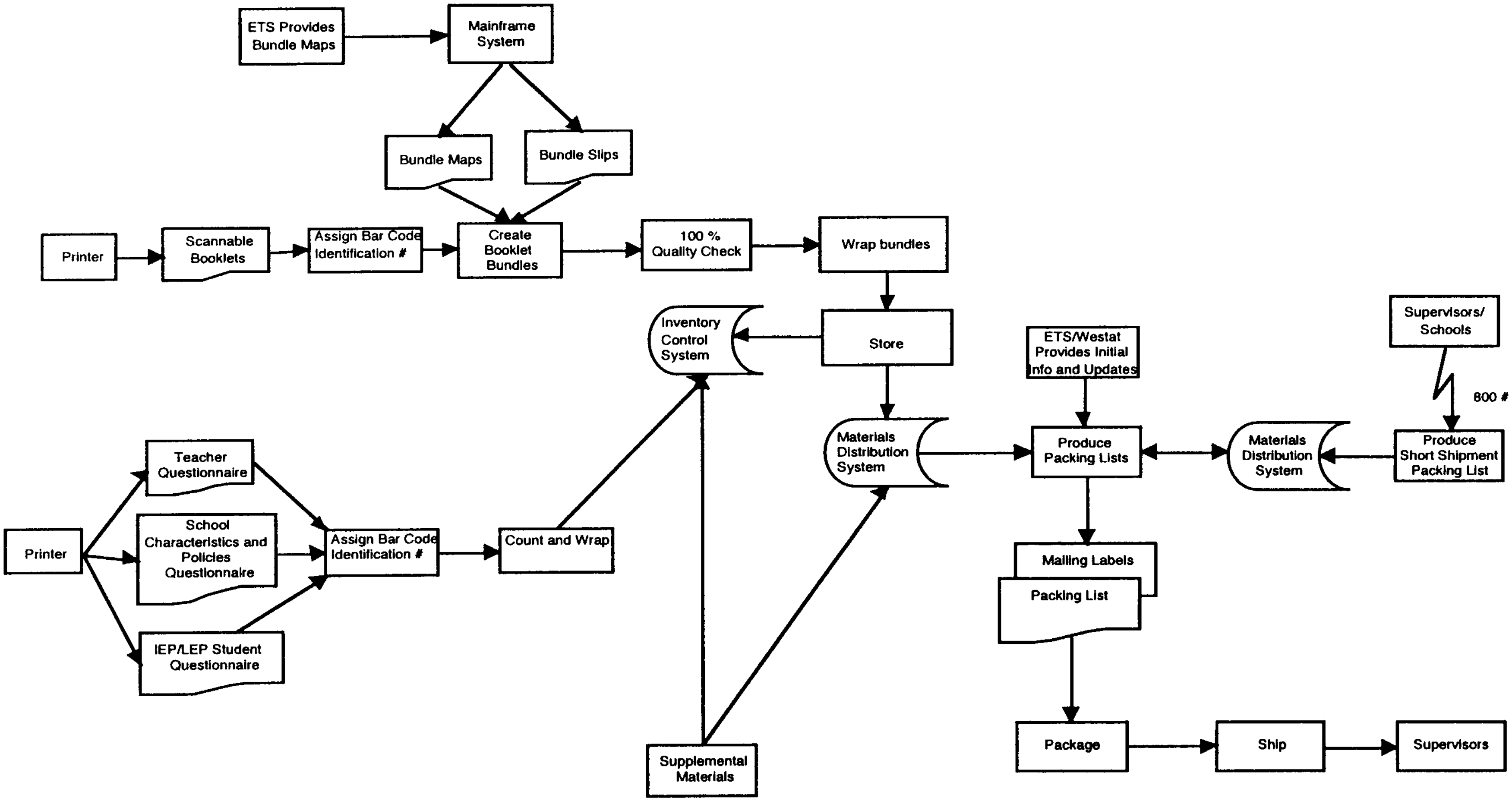
[1] The letters *D, E, F, and G* refer to the ancillary materials that accompanied the assessment booklet.

Table 5-3 (continued)
Documents Printed for the 1996 NAEP State Assessment in Science

Sample	Grade	Document (Booklet)	Subject	Type	Number of Pages	Final Copy from ETS	Approval to Print	Printed Documents Received	Quantity Printed
Main/State	4	School Questionnaire	—	Image scan	16	9/5/95	10/3/95	10/25/95	12,241
Main/State	8	School Questionnaire	—	Image scan	16	9/5/95	10/3/95	10/23/95	20,180
Main/State	8	Teacher Questionnaire	Science	OMR	24	9/21/95	10/26/95	11/1/95	46,832
Main/State	All	SD/LEP Questionnaire	—	OMR	16	10/12/95	11/9/95	12/11/95	201,050
Main/State	All	Administration Schedule	—	ICR	2	n/a	9/29/95	10/23/95	250,500
Main/State	All	SD/LEP Questionnaire Roster	—	ICR	2	n/a	10/4/95	10/25/95	101,000
Main/State	8	Teacher Questionnaire Roster	—	ICR	2	n/a	9/21/95	10/16/95	50,500

Figure 5-1

1996 NAEP State Assessment Materials Distribution Flow



The unique bundle number on the header sheet informed the system program what type of bundle should follow. A computer job was run to compare the bundle type expected to the sequence of booklets that was scanned after the header. This job also verified that the appropriate number of booklets was included in each bundle. Any discrepancies were printed on an error listing. The NCS packaging department corrected the error and the bundle was again read into the system. This process was repeated until no discrepancies existed. By using this quality-control plan, NCS could verify the document order of each bundle and account for all booklets.

Once a bundle cleared the bundle quality control process, it was shrink-wrapped and flagged on the system as ready for distribution. In the State Assessment, the bundles were not to be opened until 45 minutes before the assessment. The science bundles were shrink-wrapped, strapped, and a label was placed on the top of each bundle that read “Do Not Open Until 45 Minutes Before Assessment.”

Once all bundles for a subject area passed the bundle quality control process, information from the bundle quality control file was uploaded to the mainframe computer system and used in the creation of Administration Schedules. All Administration Schedules for each scheduled session were pre-printed with the booklet IDs designated for that session. Three bundles of booklets were pre-assigned to each session, giving each session 33 booklets. This number most closely approximated the average projected session size plus an additional supply of booklets for any extra students.

Using sampling files provided by Westat, NCS assigned bundles to schools and customized the packing lists. File data from Westat was coupled with the file of bundle numbers and the corresponding booklet numbers. This file was then used to pre-print all booklet identification numbers, school name, school number and session type, directly onto the scannable Administration Schedule. As a result, every pre-scheduled session had specific bundles assigned to it in advance. This increased the quality level of the booklet accountability system by enabling NCS to identify where any booklet should be at any time during the assessments. It also eliminated the possibility of transcription errors by assessment administrators for booklet ID numbers. Lastly, by pre-printing booklet ID numbers, the burden on the schools for transcription of data was notably reduced. NCS distributed the pre-printed Administration Schedules to state supervisors. The supervisors subsequently forwarded them to the assessment administrators in the schools before their session materials arrived. Having the preprinted Administration Schedules early assisted with sampling in the schools.

Distribution of materials for the State Assessment was accomplished in five waves or shipment dates. Except for wave “zero,” session materials were sent to a school two weeks before the assessment date. All school materials were sent directly to an assessment administrator at a school or school district. Materials for Alaska, Guam, Hawaii and Department of Defense Education Activity (DoDEA) schools were sent first in wave “zero.” These shipments required using an alternate carrier to ensure timely delivery and minimize the impact of customs delays. NCS received customs forms provided by the carrier. These forms were attached to the outside of the shipment boxes. Information such as address, school number and return address were pre-printed on these forms. Extra forms were also sent for returning boxes back to NCS in Iowa City. The remaining four waves were sent out weekly based on the schools’ scheduled assessment dates. In case any of the quantities were insufficient for the assessment, administrators were given the NAEP toll-free number to request additional materials.

Initially, a total of 9,950 sets of session materials were shipped for the 1996 State Assessments. Approximately 3,000 additional shipments of booklets and miscellaneous materials were sent. All outbound shipments were recorded in the NCS outbound mail management system. This was accomplished by having a bar code containing the school number on each address label. This bar code was read into the system, which determined the routing of the shipment and the charges. Information was recorded in a file on the system that, at the end of each day, was transferred by a PC upload to the mainframe. A computer program could then access information to produce reports on all shipments sent, regardless of the carrier used. These reports helped NCS phone staff trace shipments for state supervisors and assessment administrators.

5.3.2 Short Shipment and Phones

A toll-free telephone line was maintained for school administrators to request additional materials for the State Assessments. To process a shipment, NCS phone staff asked the caller for information such as PSU, school ID, assessment type, city, state, and ZIP code. This information was then entered into the online short shipment system and the school's mailing address was displayed on the screen to verify with the caller. The system allowed NCS staff to change the shipping address for individual requests. The clerk proceeded to the next screen that displayed the materials to be selected. After the requested items, due date and method of shipment were entered, the system produced a packing list and mailing labels. Phone staff also took phone calls concerning initial shipment delivery dates, tracing a shipment, and questions concerning NAEP. Approximately 3,750 calls were received regarding the 1996 NAEP State Assessments. Table 5-4 lists the types of requests and number of calls per request.

Table 5-4
1996 NAEP State Assessment
Phone Request Summary

Number of Calls	Request
46	Additional test booklets—increase in session
977	Additional SD/LEP questionnaires
940	Additional teacher questionnaires
515	Miscellaneous materials (excluding science kits)
212	Science kits
248	Missing materials in shipments
51	Add on school
236	Tracing shipments
400	Other (delivery dates, NAEP questions)

5.4 PROCESSING

5.4.1 Overview

The following describes the various stages of work involved in receiving and processing the documents used in the 1996 NAEP State Assessment. NCS staff created a set of predetermined rules and specifications for the processing departments within NCS to follow. Project staff performed a variety of procedures on materials received from the assessment administrators before releasing these materials into the NCS NAEP processing system. Control systems were used to monitor all NAEP materials returned from the field. The NAEP Process Control System contained the status of sampled schools for all sessions and their scheduled assessment dates. As materials were returned, the Process Control System was updated to indicate receipt dates, to record counts of materials returned, and to document any problems discovered in the shipments. As documents were processed, the system was updated to reflect processed counts. NCS report programs were utilized to allow ETS, Westat, and NCS staff to monitor the progress in the receipt control operations. The processing flow is illustrated in Figure 5-2.

An "alert" process was used to record, monitor, and categorize all discrepant or problematic situations. Throughout the processing cycle, alert situations were either flagged by computer programs or identified during clerical check-in procedures. Certain alerts, such as missing demographic information on the Administration Schedule, were resolved by opening staff retrieving the information from booklet covers.

Alert situations that could not be resolved by opening personnel were described on alert forms that were forwarded to project personnel for resolution. Once resolved, the problems and resolutions were recorded online in the Process Control System.

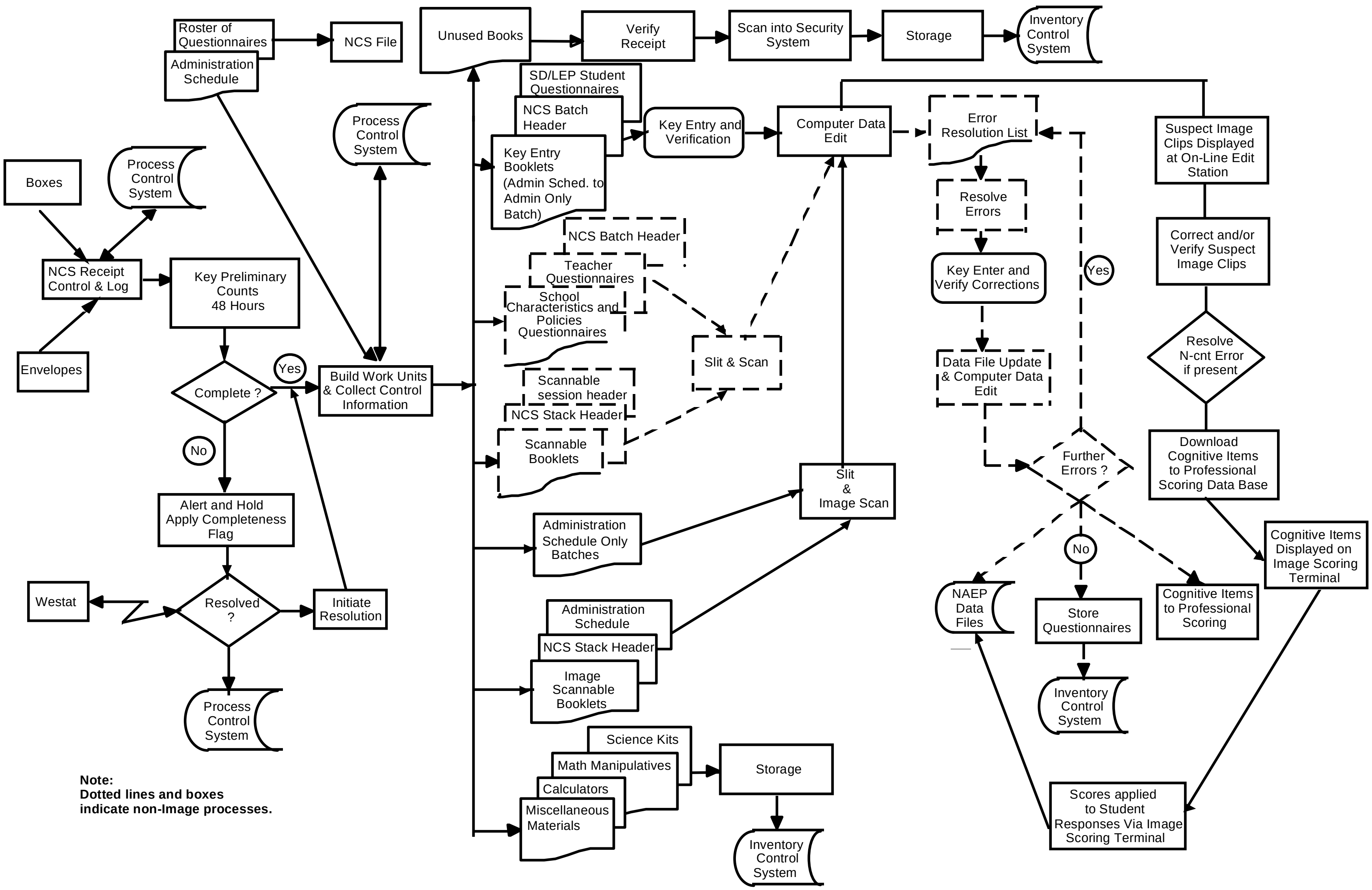
NCS's Workflow Management System was used to track batches of student booklets through each processing step, allowing project staff to monitor the status of all work in progress. It was also used by NCS to analyze the current work load, by project, across all work stations. By routinely monitoring these data, NCS's management staff was able to assign priorities to various components of the work and to monitor all phases of the data receipt and processing.

5.4.2 Document Receipt and Tracking

All shipments were to be returned to NCS packaged in their original boxes. As mentioned earlier, NCS packaging staff applied a bar code label to each box indicating the NAEP school ID number. When a shipment arrived at the NCS dock area, this bar code was scanned to a personal computer file, and the shipment was forwarded to the receiving area. The personal computer file was then transferred to the mainframe and the shipment receipt date was applied to the appropriate school within the Process Control System, providing the status of receipts regardless of any processing delays. Each receipt was reflected on the Process Control System status report provided to the NCS receiving department and supplied to Westat via electronic file transfer and in hard-copy format. ETS also received a hard copy. The Process Control System file could be manually updated to reflect changes, if necessary.

Figure 5-2

1996 NAEP State Assessment Materials Processing Flow Chart



Note:
Dotted lines and boxes indicate non-Image processes.

Receiving personnel also checked the shipment to verify that the contents of the box matched the school and session indicated on the label. Each shipment was checked for completeness and accuracy. Any shipment not received within two days of the scheduled assessment date was flagged in the Process Control System and annotated on the Process Control System report. The administration status of these delayed shipments was checked and in some cases a trace was initiated on the shipment.

A new requirement for NCS was to open all shipments within 48 hours of their receipt and to key-enter preliminary processing information into the Process Control System from the Administration Schedule. The preliminary information was written on the Administration Schedule by assessment administrators and consisted of the following:

- School number
- Session number
- Original test date
- Total number to be assessed
- Total number assessed

This preliminary information, used to provide Westat with timely student response rates, was updated with actual data when materials passed through processing error free. A completeness flag was also applied to the process control file by NCS opening staff if any part of the shipment was missing.

If multiple sessions were returned in one box, the contents of the package were separated by session. The shipment was checked to verify that all booklets preprinted or handwritten on the Administration Schedule were returned with the shipment and that all administration codes from booklet covers matched the Administration Schedule. If discrepancies were discovered at any step in this process, the receiving staff issued an alert to facilitate tracking.

If the administrator indicated that a make-up session was being held the documents were placed on holding carts until the make-up session documents arrived. If no make-up session was indicated, Westat was contacted for the status of the missing materials. If the missing materials were to be returned, the documents already received were held until that time. If the materials were not being returned, processing continued and the appropriate administration code was applied to the Administration Schedule.

Once all booklets listed on the Administration Schedule for a session were verified as present, the entire session (both the Administration Schedule and booklets) was batched by grade level and session type. Each batch was assigned a unique batch number. This number, created on the Image Capture Environment system for all image-scannable documents and on the Workflow Management System for all key-entry and OMR-scannable documents, facilitated the internal tracking of the batches and allowed departmental resource planning. All other scannable documents (School Characteristics and Policies Questionnaires, Teacher Questionnaires, SD/LEP Questionnaires, and rosters) were batched by document type in the same manner.

Because the State Assessment science booklets were image-scannable, batch numbers for these documents were created on the Image Capture Environment System. Sessions were sorted by grade level and automatically uploaded to the Workflow Management System after batch creation. The Administration Schedule for these document types was used as a session header within a batch.

The 1996 NAEP State Assessment utilized one roster to document and track the School Characteristics and Policies Questionnaire and the Students with Disabilities/Limited English Proficiency (SD/LEP) Questionnaire. In addition, the State Assessment used the Teacher Questionnaire Roster to record the distribution and return of Teacher Questionnaires.

Some questionnaires may not have been available for return with the shipment. These were returned to NCS at a later date in an envelope provided for that purpose. The questionnaires were submitted for scanning as sufficient quantities became available for batching.

Receipt of the questionnaires was entered into the system using the same process as was used for the Administration Schedule described in previous sections. The rosters were grouped with other rosters of the same type from other sessions, and a batch was created on the Image Capture Environment system. The batch was then forwarded to scanning where all information on the rosters was scanned into the system.

In the 1996 NAEP State Assessment, NCS used a sophisticated booklet accountability system to track all distributed booklets. As stated earlier, prior to the distribution of NAEP materials, unique booklet numbers were read by bundle into a file. Specific bundles were then assigned to particular supervisors or schools. This assignment was recorded in the NAEP Materials Distribution System. When shipments arrived at NCS from the field, all used booklets were submitted for processing and a "processed documents" file was maintained. Unused booklets were submitted for security scanning where booklet ID bar codes were read and recorded into a separate file. This file and the "processed documents" file were later compared to the original bundle security file for individual booklet matching. A list of unmatched booklet IDs was printed in a report used to confirm non-receipt of individual booklets. Efforts were made to be sure unused materials from the State Assessment were returned by school personnel. The used but returned booklet IDs were also read by the bar code scanner and added to the bundle security file. All unused materials received were then inventoried and sent to the NCS warehouse for storage while awaiting authorization from ETS to salvage them.

The transcription of the student response data into machine-readable form was achieved through the use of the following three separate systems: data entry (which included optical mark recognition (OMR) and image scanning, ICR, and key entry), data validation (edit), and data resolution.

5.4.3 Data Entry

The data entry process was the first point at which booklet-level data were directly available to the computer system. Depending on the NAEP document, one of three methods was used to transcribe NAEP data to a computerized form. The data on scannable documents were collected using NCS optical-scanning equipment that also captured images of the constructed-response items and ICR fields. Nonscannable materials were keyed through an interactive online system. In both of these cases, the data were edited and suspect cases were resolved before further processing.

All student booklets, questionnaires, and control documents were scannable. Throughout all phases of processing, the student booklets were batched by grade and session type. The scannable documents were then transported to a slitting area where the folded and stapled spine was removed from the document. This process utilized an “intelligent slitter” to prevent slitting the wrong side of the document. The documents were jogged by machine so that the registration edges of the NAEP documents were smoothly aligned, and the stacks were then returned to the cart to be scanned.

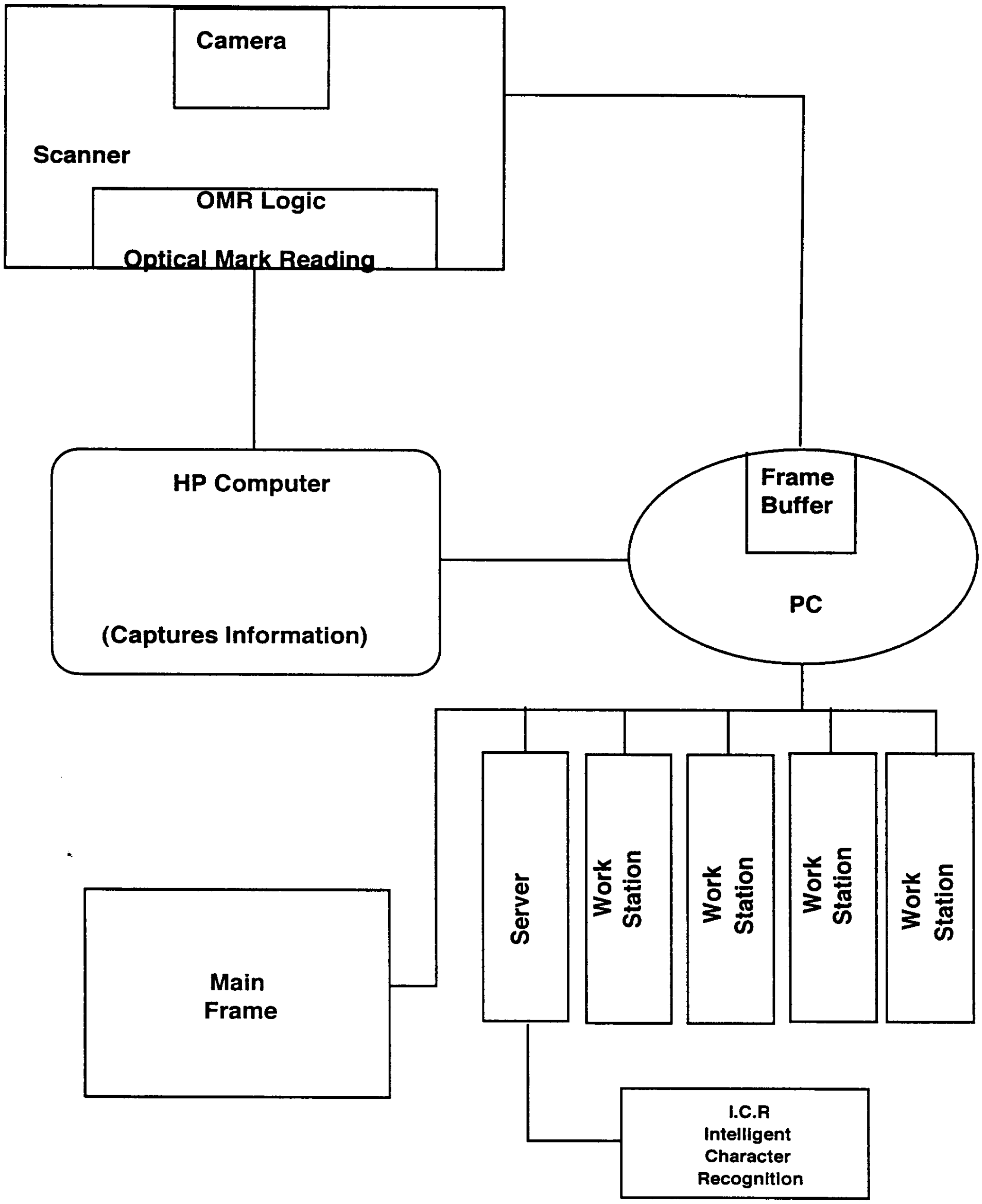
During the scanning process (shown in Figure 5-3), each scannable NAEP document was uniquely identified using a print-after-scan number consisting of the scan batch number, the sequential number within the batch, and the bar code ID of the booklet. These numbers were printed on each sheet of each document as it exited the scanner. This permitted the data editors to quickly and accurately locate specific documents during the editing phase. The print-after-scan number remained with the data record, providing a method for easy identification and quick retrieval of any document.

The data values were captured from the booklet covers and Administration Schedules and were coded as numeric data. Unmarked fields were coded as blanks and editing staff were alerted to missing or uncoded critical data. Fields that had multiple marks were coded as asterisks (*). The data values for the item responses and scores were returned as numeric codes. The multiple-choice single response format items were assigned codes depending on the position of the response alternative; that is, the first choice was assigned the code “1,” the second “2,” and so forth. The mark-all-that-apply items were given as many data fields as response alternatives; the marked choices were coded as “1” while the unmarked choices were recorded as blanks. The images of constructed-response items were saved as a digitized computer file. The area of the page that needed to be clipped was defined prior to scanning through the document definition process. The fields from unreadable pages were coded “X” as a flag for resolution staff to correct. In addition to capturing the student responses, the bar code identification numbers used to maintain process control were decoded and transcribed to the NAEP computerized data file.

As the scanning program completed scanning each stack, the stack was removed from the output hopper and placed in the same order they were scanned on the output cart. The next stack was removed from the input cart and placed into the input hopper, after which the scanning resumed. When the operator had completed processing the last stack of the batch, the program was terminated. This closed the dataset that automatically became available for the data validation (edit) process. The scanned documents were then forwarded to a holding area in case they needed to be retrieved for resolution of edit errors.

NCS again used the ICR engine to read various hand and machine printing on the front cover of the assessment and supervisor documents for the 1996 NAEP assessments. Some information from scannable student documents, such as the Administration Schedule, the Roster of Questionnaires, and some questions in the School Characteristics and Policies Questionnaires, were read by the ICR engine and verified by an online key-entry operator. In all, the ICR engine read approximately 15 million characters. The ICR engine saved NAEP field staff and school personnel a significant amount of time because they no longer had to enter this data by gridding rows and columns of data.

Figure 5-3
1996 NAEP State Assessment
Image Scanning Flow Chart



NCS also implemented new programs that allowed the scanners to read imprinted codes, known as 2-out-of-5 codes, that were printed via a Xerox 4280 printer on the Administration Schedule. These 2-out-of-5 codes were imprinted at the same time the booklet ID numbers were printed on the Administration Schedule and identified which booklet IDs were listed on that document. When the scanning programs were unable to translate the 2-out-of-5 codes (thereby identifying the booklet ID numbers on the document) image clips of the booklet ID numbers were displayed to online editing staff for verification. This eliminated a significant amount of online editing time needed to process the NAEP assessments.

To provide another quality check on the image scanning and scoring system, NCS staff stamped blank booklets with a rubber stamp and assigned these booklets mock scores from the valid range. Each unique item type scored via the image system had two quality control stamps per valid score. An example of the stamp used is given below.

The quality control booklets were batched and processed together with student documents of the same type. Because all of a specific item were batched together for transmission to the scoring facility, the quality control-stamped responses were integrated with the student responses and transmitted simultaneously to the scoring facility. During the scoring process, both student responses and the quality control items were randomly displayed so scores could be applied.

When a person who was scoring responses (reader) later saw the quality control sample on the monitor during scoring, he or she was to notify the team leader, who confirmed the score assigned by the reader was the score listed on the sample. The quality control booklets were included in the pool of all items to be drawn from for the 25 percent reliability rescore.

All image quality-assurance documents were created prior to the beginning of scoring and all pre-determined score points were used. Because during the process of scoring, valid score points can be changed or dropped completely, NCS provided ETS with documentation explaining what quality control documents were produced and which score points on these items were no longer valid. When an image quality control stamp was displayed to a reader that contained a score point that was no longer valid, the reader gave the response a score point of zero.

A key entry and verification process was used to make corrections to the teacher questionnaires and the SD/LEP student questionnaires. The Falcon system that was used to enter these data is an online data entry system designed to replace most methods of data input such as keypunch, key-to-disk, and many of the microcomputer data entry systems. The terminal screens were uniquely designed for NAEP to facilitate operator speed and convenience. The fields to be entered were titled to reflect the actual source document.

5.4.4 Data Validation

Each dataset produced by the scanning system contains data for a particular batch. These data had to be validated (or edited) for type and range of response. The data-entry and resolution system used was able to simultaneously process a variety of materials from all age groups, subject areas, control documents, and questionnaires as the materials were submitted to the system from scannable and non-scannable media.

The data records in the scan file were organized in the same order in which the paper materials were processed by the scanner. A record for each batch header preceded all data records for that batch. The document code field on each record distinguished the header record from the data records.

When a batch-header record was read, a pre-edit data file and an edit log were generated. As the program processed each record within a batch from the scan file, it wrote the edited and reformatted data records to the pre-edit file and recorded all errors on the edit log. The data fields on an edit log record identified each data problem by the batch sequence number, booklet serial number, section or block code, field name or item number, and data value. After each batch had been processed, the program generated a listing or online edit file of the data problems and resolution guidelines. An edit log listing was printed at the termination of the program for all non-image documents. Image “clips” requiring editing were routed to online editing stations for those documents that were image scanned.

As the program processed each data record, it first read the booklet number and checked it against the session code for appropriate session type. Any mismatch was recorded on the error log and processing continued. The booklet number was then compared against the first three digits of the student identification number. If they did not match, a message was written on the error log. The remaining booklet cover fields were read and validated for the correct range of values. The school codes had to be identical to those on the Process Control System record. All data values that were out of range were read “as is” but were flagged as suspect. All data fields that were read as asterisks (*) were recorded on the edit log or online edit file.

Document definition files described each document as a series of blocks that in turn were described as a series of items. The blocks in a document were transcribed in the order that they appeared in the document. Each block’s fields were validated during this process. If a document contained suspect fields, the cover information was recorded on the edit log along with a description of the suspect data. The edited booklet cover was transferred to an output buffer area within the program. As the program processed each block of data from the dataset record, it appended the edited data fields to the data already in this buffer.

The program then cycled through the data area corresponding to the item blocks. The task of translating, validating, and reporting errors for each data field in each block was performed by a routine that required only the block identification code and the string of input data. This routine had access to a block definition file that had, for each block, the number of fields to be processed, and, for each field, the field type (alphabetic or numeric), the field width in the data record, and the valid range of values. The routine then processed each field in sequence order, performing the necessary translation, validation, and reporting tasks.

The first of these tasks checked for the presence of blanks or asterisks (*) in a critical field. These were recorded on the edit log or online edit file and processing continued with the next field. No action was taken on blank fields for multiple-choice items because the asterisk code indicated a non-response. The field was validated for range of response, and any values outside of the specified range were recorded on the edit log or online edit file. The program used the item-type code to make a further distinction among constructed-response item scores and other numeric data fields.

Moving the translated and edited data field into the output buffer was the last task performed in this phase of processing. When the entire document was processed, the completed string of data was written to the data file. When the program encountered the end of a file, it closed the dataset and generated an edit listing for non-image and key-entered documents. Image-scanned items that required correction were displayed at an online editing terminal.

5.4.5 Editing for Non-Image and Key-Entered Documents

Throughout the system, quality procedures and software ensured that the NAEP data were correct. All student documents on the Administration Schedule were accounted for, as receipt control personnel checked that the materials were undamaged and assembled correctly. The machine edits performed during data capture verified that each sheet of each document was present and that each field had an appropriate value. All batches entered into the system, whether key-entered or machine-scanned, were edited for errors.

Data editing took place after these checks. This consisted of a computerized edit review of each respondent's document and the clerical edits necessary to make corrections based upon the computer edit. This data-editing step was repeated until all data were correct.

The first phase of data editing was designed to validate the population and ensure that all documents were present. A computerized edit list, produced after NAEP documents were scanned or key entered, and all the supporting documentation sent from the field were used to perform the edit function. The hard-copy edit list contained all the vital statistics about the batch: number of students, school code, type of document, assessment code, suspect cases, and record serial numbers. Using these inputs, the data editor verified that the batch had been assembled correctly and that each school number was correct.

During data entry, counts of processed documents were generated by type. These counts were compared against the information captured from the Administration Schedules. The number of assessed and absent students processed had to match the numbers indicated on the Process Control System.

In the second phase of data editing, experienced editing staff used a predetermined set of specifications to review the field errors and record necessary corrections to the student data file. The same computerized edit list used in phase one was used to perform this function. The editing staff reviewed the computer-generated edit log and the area of the source document that was noted as being suspect or as containing possible errors. The composition of the field was shown in the edit box. The editing staff checked this piece of information against the NAEP source document. At that point, one of the following took place:

Correctable error. If the error was correctable by the editing staff as per the editing specifications, the correction was noted on the edit log for later correction via key-entry.

Alert. If an error was not correctable as per the specifications, an alert was issued to NAEP project staff for resolution. Once the correct information was obtained, the correction was noted on the edit log for key-entry correction.

Non-correctable error. If a suspected error was found to be correct as stated and no alteration was possible according to the source document and specifications, the programs were tailored to allow this information to be accepted into the data record. No corrective action was taken.

The corrected edit log was then forwarded to the key-entry staff for processing. When all corrections were entered and verified for a batch, an extract program pulled the corrected records into a mainframe dataset. At this point, the mainframe edit program was initiated. The edit criteria were again applied to all records. If there were further errors, a new edit listing was printed and the cycle was repeated.

When the edit process produced an error-free file, the booklet ID number was posted to the NAEP tracking file by age, assessment, and school. This permitted NCS staff to monitor the NAEP processing effort by accurately measuring the number of documents processed by form. The posting of booklet IDs also ensured that a booklet ID was not processed more than once.

5.4.6 Data Validation and Editing of Image-Processed Documents

The paper edit log for key-entered documents was replaced by online viewing of suspect data for all image-processed documents. For rapid resolution, the edit criteria for each item in question appeared on the screen along with the suspect item. Corrections were made immediately. The system employed an edit/verify system that ultimately meant that two different people viewed the same suspect data and operated on it separately. The “verifier” made sure the two responses (one from either the entry operator or the ICR engine) were the same before the system accepted that item as being correct. The verifiers could either overrule or agree with the original correction made if the two did not match. If the editor could not determine the appropriate response, he or she escalated the suspect situation to a supervisor. For errors or suspect information that could not be resolved by supervisory staff, a product-line queue was created. This allowed supervisors to escalate edits to project staff for resolution. By having this product-line queue, project staff were able to quickly locate edit clips within the image system, speeding up the resolution process.

Once an entire batch was through the edit phase, it became eligible for the count-verification phase. The Administration Schedule data were examined systematically for booklet IDs that should have been processed (assessed administration codes). All documents under that Administration Schedule were then inspected to ensure that all of the booklets were included.

With the satisfactory conclusion of the count-verification phase, the edited batch file was uploaded to the mainframe, where it went through yet another edit process. A paper edit log was produced and, if errors remained, was forwarded to another editor. When this paper edit was satisfied, the Process Control System and Workflow Management System were updated. Because there was a possible time lag between a clean edit in the image system and a clean edit in the mainframe systems, the batch was not archived until 48 hours after the image edit phase was completed.

5.4.7 Data Transmission

Due to the rapid pace of scoring on an item-by-item basis, the NCS scoring specialists found it necessary to continually monitor the status of work available to the readers and plan the scoring schedule several weeks in advance. On Wednesday of each week, the NCS performance assessment specialist in charge of each subject area planned the next two weeks' schedule. That information was then provided to the person in charge of downloading data to the scoring center. By planning the scoring schedule two weeks in advance, the scoring specialists were able to ensure that readers would have sufficient work for at least one week, after which the next download would occur to supplement the volume of any unscored items and add an additional week's work to the pool of items to score. Additionally, by scheduling two weeks' data transmission, flexibility was added to the scoring schedule, making it possible to implement last-minute changes in the schedule once the items had been delivered to the scoring center. Depending on the number of items to be transmitted, the actual downloading was conducted on Friday or was divided into two smaller sessions for Thursday and Friday download. By the first week of May 1996, there was sufficient space on the scoring servers to load all remaining unscored items to the scoring center.

Delivery of data to the scoring center was accomplished via several T1 transmission lines linking the mainframe computers and the NAEP servers at the document-scanning site in the NCS main facility with the scoring servers dedicated to distributing work to the professional readers at the scoring center. The actual task of scheduling items for downloading was accomplished using a code written by the Image Software Development team. This code enabled the person scheduling the download to choose a team of readers and select the scheduled items from a list of all items that that team would be scoring throughout the scoring project. This process was repeated for all teams of readers until all anticipated work was scheduled. Once this task was completed, the scheduled job was tested to determine if there was sufficient free disk space on the servers at the scoring center. If for any reason sufficient disk space was not available, scheduled items could be deleted from the batch individually or as a group until the scheduled batch job could accommodate all items on the available disk space at the scoring center. Once it was determined that sufficient disk space was available, transmission of student responses commenced. Data transmission was typically accomplished during off-shift hours to minimize the impact on system-load capacity.

5.5 PROFESSIONAL SCORING

5.5.1 Overview

Scoring of the 1996 NAEP State Assessment constructed-response items was conducted using NCS's imaging technology. All 1996 responses were scored online by readers working at image stations. The logistical problems associated with handling large quantities of student booklets were removed for those items scored on the image system.

One of the greatest advantages image technology presented for NAEP scoring was in the area of sorting and distributing work to scorers. All student responses for a particular item, regardless of where spiraling had placed that item in the various booklet forms, were grouped together for presentation to a team of readers. This allowed training to be conducted one item at a time, rather than in blocks of related items, thus focusing readers' attention on the complexities of a single item.

A number of tools built into the system allowed table leaders and trainers to closely and continuously monitor reader performance. A detailed discussion of these tools can be found later in this chapter.

The system automatically routed six percent of student responses to other members of the team for second scoring. Readers were given no indication of whether the response had been scored by another reader, thereby making the second scoring truly blind. On-demand, real-time reports on interreader reliability (drawn from those items that were second-scored) presented extremely valuable information on team and individual scoring. Information on adjacent and perfect agreement, score distribution, and quantity of responses scored were continuously available for consultation. Similarly, back-reading of student responses could be accomplished in an efficient and timely manner. Also, table leaders were able to read a large percentage of responses, evaluating the appropriateness and accuracy of the scores assigned by readers on their teams.

Project management tools assisted table leaders in making well-informed decisions. For example, knowledge of the precise number of responses remaining to be scored for a particular item allowed table leaders to determine the least disruptive times for lunch breaks.

Both readers and table leaders responded with enthusiasm to the system, remarking on the ease with which student responses could be read and on the increased sense of professionalism they felt in working in this technological environment. Readers took periodic breaks, in addition to their lunch break, to reduce the degree of visual fatigue. Readers were grouped in teams of 9 to 14 readers per team; each team working with a specific table leader.

5.5.2 Training Paper Selection

In March 1996, a pool of papers to be used during training was selected by NCS team leaders chosen for their credentials in science. The team leaders sent the pool of papers to ETS assessment division subject specialists, who created the master training set. Team leaders were used for this task because it gave them the advantages of working on specific items, learning the make-up of the various booklets, learning the terminology, and understanding the processing of the booklets at NCS. This was especially important in 1996, because most scoring activities occurred via the image processing system.

Generally, the training set for each short (two- or three-point) item included 40 papers:

- 10 anchor papers
- 20 practice papers
- 5 papers in Calibration Set #1
- 5 papers in Calibration Set #2

Generally, the training set for each extended (four- or five-point) item included 75 papers:

- 15 anchor papers
- 40 practice papers
- 10 papers in each of two qualification sets
- 5 papers in Calibration Set #1
- 5 papers in Calibration Set #2

Anchor papers, or sets, are those papers that represent the best examples of each score point. They are used to illustrate the scoring guide so that the reader can return to this set and compare it with student responses during scoring. Practice papers, or sets, include the remainder of the scored examples, excluding the scores, so that the reader can practice on some student responses prior to scoring. The purpose is to elicit discussion and give scorers a chance to ask questions. Qualification sets are used by the trainer to ensure that each reader has understood the scoring guide and can apply it to student documents. Similar to practice papers, the scores are masked so the reader can assign a score. A predetermined number of scores must be correct for the reader to remain on the scoring project. Calibration sets are used after a long break in scoring has occurred (e.g., after lunch in the early days of a project, or first thing in the morning) to ensure that the readers review the scoring guide and the anchor papers, and to prevent the scorers from drifting to the middle range of possible score points.

To ensure that the ETS assessment specialist would have a wide range of student responses to encompass all score points, NCS personnel copied approximately 125 papers for each five-point item, 100 papers for each four-point item, 75 papers for each three-point item, and 50 papers for each two-point item. To ensure that training papers represented the range of responses obtained from the sample population, NCS personnel selected papers randomly from across the sample. The student identifier (barcode) was written on the copy and NCS team leaders assigned tentative scores to the responses. The responses were numbered sequentially, copied, and sent via overnight delivery to ETS. When the training packet was compiled, the ETS assessment specialist faxed the composition of the packets back using the sequential numbers. ETS staff kept its copy of the training sets.

From the faxed sheets, packets were created for each item using the original copies of the student responses. These packets were then forwarded to the NCS communication center for copying, and stored for the teams' use in training. ETS also sent the most up-to-date version of the training packet to the NCS scoring center for each item to be included in the scoring guide.

5.5.3 General Training Guidelines

ETS and NCS personnel conducted training for the constructed-response items on an item-by-item basis, so that each item could be scored immediately after training. In all, 24 table leaders and 306 readers worked from March 18 to June 7, 1996, to complete scoring for the 1996 NAEP State Assessments. Each member of a team received a copy of the stimulus and training materials for the items that his or her team would be scoring. Before training, each team member became familiar with those materials under the guidance of the trainer who explained the anchor papers, exemplifying the various score-point levels. Next, ETS and NCS staff (the trainers) conducted training sessions to explain the anchor papers, exemplifying the various score point levels. The team proceeded with each member scoring the practice papers, and then discussing those papers as a group while the trainer clarified issues and answered questions. The papers selected for each training set were chosen to illustrate a range from easily classifiable responses to borderline responses for each score point.

When the trainer was confident the readers were ready to begin scoring short constructed responses, the table leader signaled the system to release the responses to the team members who had successfully completed training. For extended constructed-response items, each team member was given a qualifying set that had been prescored by the trainer in conjunction with the table leader. Readers were required to score an exact match on 80 percent of the items in order to qualify for scoring. If a reader failed on the first attempt, the trainer discussed the discrepant scores with the reader and administered a second qualifying set. Again, 80 percent exact agreement was required to score the item. During the beginning stages of scoring, the team members discussed student responses with the trainer and table leader to ensure that issues not addressed in training were handled in the same manner by all team members.

After the initial training, readers scored the items, addressing questions to the table leader and/or trainer when appropriate. Depending upon the number of responses, length of responses and complexity of the rubric, scoring of an individual item ranged anywhere from one-half hour to two weeks. Whenever a break longer than 15 minutes occurred in scoring, each team member received a set of calibration papers that had been prescored by the trainer and table leader. Each team member scored the calibration set individually, and then the team discussed the papers to ensure against scorer drift.

5.5.4 Table Leader Utilities and Reliability Reports

Among the many advantages of the image scoring system is the ease with which workflow to readers can be regulated and scoring can be monitored. After training, the table leader would route work only to those scorers qualified for a particular item. He or she could also cancel a reader’s qualification to score an item if review of a reader’s work indicated inaccurate scoring and that supplemental training was necessary after scoring had begun.

After scoring began, NCS table leaders reviewed each reader’s progress using a backreading utility that allowed the table leader to review papers scored by each reader on the team. Typically, a table leader reviewed responses scored by each reader at the same rate at which second scoring occurred (i.e., six percent for items with both state and national samples and more for items with only a national sample). Table leaders made certain to note the score the reader awarded each response as well as the score a second reader gave that same paper. This was done as an interreader reliability check. Reliability information for the current assessment is in Section 5.5.5.3. Alternatively, a table leader could choose to review all responses given a particular score to determine if the team as a whole was scoring consistently. Both of these review methods used the same display screen and showed the ID number of the reader and the scores awarded. If the table leader disagreed with the score given an item, he or she discussed it with the reader for possible correction. Replacement of scores by the table leader was done only with the knowledge and approval of the reader, thereby serving as a learning experience for the reader. In the case where the response was second scored, neither score was changed.

The table leaders were able to monitor workflow using a status tool that displayed the number of items scored, the number of items first-scored that still needed to be second-scored, the number of items remaining to be second-scored, and the total number of items remaining to be scored. This allowed the team leaders and performance assessment specialists to accurately monitor the rate of scoring and to estimate the time needed for completion of the various phases of scoring.

5.5.5 Main and State NAEP Science

The science portion of the 1996 NAEP included a total of 374 discrete constructed-response items (see Table 5-5). It was scored over three segments and two shifts (see Table 5-6). Many kinds of constructed-response items were utilized in the assessment to measure different elements of students' conceptual understanding of scientific material as well as their practical reasoning ability. The items scored included short-answer constructed responses and extended constructed responses. Each constructed-response item had a unique scoring guide that identified the range of possible scores for the item and defined the criteria to be used in evaluating student responses.

During the course of the project, each team scored short constructed-response items using a scale that allowed for partial credit as follows:

"1" = incorrect response
"2" = partial understanding
"3" = correct response

The readers also scored extended constructed-response items on a scale of "1" to "4" as follows:

"1" = incorrect response
"2" = minimal understanding
"3" = satisfactory level of comprehension
"4" = correct reasoning

Table 5-5
Number of Constructed Response Items
for Science State and National Assessments

Grade	2-Point Items	3-Point Items	4-Point Items	5-Point Items	6-Point Items	Total
4	0	58	10	3	0	71
4/8	0	51	8	1	0	60
8	0	46	13	0	0	59
8/12	0	54	6	0	0	60
12	0	83	34	7	1	125
Total	**0**	**292**	**71**	**11**	**1**	**375**

Table 5-6
1996 NAEP Science State and National Assessments
Readers and Dates

Segment	Number of Table Leaders	Number of Scorers	Dates
Segment 1 Days	3	24	3/18/96-4/5/96
Segment 2 Days	3	27	4/8/96-5/3/96
Segment 2 Evenings	5	70	4/8/96-5/2/96
Segment 3 Days	12	156	5/6/96-6/7/96
Segment 3 Evenings	12	150	5/6/96-6/7/96

5.5.5.1 Science Training

The training on each item was conducted by science specialists from ETS and NCS. The first teams began training on March 18, 1996. Other teams were phased in throughout the project. Hands-on items were scored a block at a time with a unique scoring guide for each item because of the related nature of the items. The rest of the assessment was scored item-by-item so that each reader worked on only one set of rubrics at a time. After scoring all available responses, a team would then proceed with training and scoring the next item. Scoring was completed on June 7, 1996. Table 5-6 provides detailed information on the dates of scoring and the number of readers and table leaders.

Training involved explaining the item and its scoring guide to the team and discussing responses that represented the various score points in the guide. Typically, two or three anchor responses were chosen for each score point. During this stage, readers and the table leader kept notes of scoring decisions. The table leader was then responsible for compiling those notes and ensuring that all readers were in alignment. When review of the anchor packet was completed, the readers scored and discussed 10 to 20 prescored “practice papers” that represented the entire range of score points the item could receive. After the trainer and table leader determined that the team had reached consensus, the table leader then released work on the image-scoring system to the readers. The readers would initially take turns reading their first “live” responses to the team or work in pairs as a final check before beginning work individually. Once the practice session was completed, the formal scoring process began.

5.5.5.2 Science Scoring

All scoring for science was conducted by way of an image-based scoring system. During scoring, the team leaders continued to compile notes on scoring decisions for the readers' reference and guidance. Additionally, table leaders closely monitored interreader reliability using both team and individual statistics as a reference. Consistently throughout the scoring of each item, the table leaders also performed backreading duties in which they reviewed a sample of the responses scored by each reader on the team. Lead scorers selected for their experience and accuracy in scoring assisted the table leaders in backreading. The table leaders and performance assessment specialist continuously monitored the progress of each team and noted all scoring-related decisions to ensure that training and scoring progressed smoothly and in a timely manner.

One advantage of utilizing an image-based scoring system is the ability to construct reader aids to simplify scoring, thus increasing reader reliability. Prior to the start of the project, the ETS subject area specialist and the NCS performance assessment specialist identified several items for the construction of overlays. Overlays serve as templates to define boundaries in which correct responses must be located or allow the placement of correct answers directly on the displayed image, and are displayed along with the student response. A schematic representation of each overlay was included with the scoring guide and sample papers for these items to familiarize readers with the use of the scoring aids during training.

In the 1996 State Assessment in science, 3,926 booklets were processed and 92,565 constructed response items were scored for the fourth-grade sample, and 113,065 booklets were processed and 3,000,014 constructed response items were scored for the eighth-grade sample.[2] Table 5-7 provides more detailed information by grade and book type (spiral and advanced) for both the national and state assessments in science.

Table 5-8 shows the codes that were used for unscorable science items.

[2] The 1996 State Assessment in science was conducted at grade 8 only, except for Department of Defense Education Activity (DoDEA) schools that were assessed at both grades 4 and 8.

Table 5-7
1996 Science State and National Assessments
Constructed-Response Items Scored

Science Grade	Data	Type: Regular	Hands-on	Advanced	Grand Total	Assessment Proportions: National	State
4	Unique Items	70	24	0	94	74.8%	25.2%
	Responses First Scored	200,319	94,004	0	294,323	220,271	74,052
	Responses Second Scored	50,080	23,501	0	73,581	55,068	18,513
	First and Second Scored	250,399	117,505	0	367,904	275,339	92,565
	Average Percentage Agreement	93.9	93.8	N/A	93.9		
8	Unique Items	94	31	0	125	9.7%	90.3%
	Responses First Scored	2,157,377	976,844	0	3,134,222	304,020	2,830,202
	Responses Second Scored	129,443	58,611	0	188,053	18,241	169,812
	First and Second Scored	2,286,820	1,035,455	0	3,322,275	322,261	3,000,014
	Average Percentage Agreement	93.4	95.0	N/A	93.8		
12	Unique Items	94	26	36	156		
	Responses First Scored	198,563	75,120	88,166	361,849		
	Responses Second Scored	49,641	18,780	22,041	90,462		
	First and Second Scored	248,204	93,900	110,207	452,311		
	Average Percentage Agreement	93.0	94.4	94.5	93.6		
Total Unique Items		258	81	36	375		
Total Responses First Scored		2,556,260	1,145,968	88,166	3,790,394		
Total Responses Second Scored		229,163	100,892	22,041	352,096		
Total First and Second Scored		2,785,423	1,246,860	110,207	4,142,490		
Total Average Percentage Agreement		93.4	94.4	94.5	93.7		

Table 5-8
1995-1996 NAEP Assessments
Codes For Unscorable Science Items

Code	Type of Response
0	Blank or random marks
8	Completely crossed-out or erased
9	"I don't know," refusal, off-task, illegible, or language other than English that could not be translated

5.5.5.3 Reliability of Scoring Science Constructed-Response Items

A minimum of 25 percent of the responses for science items involved only in the national assessment (grade 12 items) and six percent of the responses for science items involved in both the national and state assessments (grade 4 and grade 8 items) were scored by a second reader to obtain statistics on interreader (interrater) reliability. Ranges for percentage of exact agreement for state and national assessments of science, together, can be found in Table 5-9. Average percentage of exact agreement for each booklet type (spiral and advanced) can be found in Table 5-7. This reliability information was also used by the team leaders to monitor the

Table 5-9
1996 Science State and National Assessments
Ranges of Percentage Agreement Among Readers

Assessment	Number of Unique Items Total	Number of Items in Percentage Agreement Range 60-69%	70-79%	80-89%	Above 90%
4th grade	94	0	0	13	81
8th grade	125	0	0	20	105
12th grade	156	0	0	26	130

capabilities of all readers and maintain uniformity of scoring across readers. Reports of reliability information could be generated on demand by the table leader, team leader, or performance assessment specialist when needed, and they were displayed at a computer workstation. In addition to the immediate feedback provided by the online reliability reports, each table leader could also review the actual responses scored by a reader by using the backreading tool. In this way, the table leader was able to monitor each reader carefully and correct difficulties in scoring almost immediately with a high degree of efficiency.

In addition to reliability information calculated and used during the scoring process, several additional reliability measures are calculated for constructed-response items after the item response data has been placed on the NAEP database. These include a final percentage exact agreement, the intraclass correlation, Cohen's kappa, and the product-moment correlation between the scores for the first and second readers. These measures are summarized in Zwick (1988), Kaplan and Johnson (1992), and Abedi (1996). Each measure has advantages and disadvantages for use in different situations. In *The NAEP 1996 Technical Report* (Allen, Carlson, & Zelenak, 1998), the percentage exact agreement is reported for all constructed-response items, Cohen's kappa is reported for dichotomously scored constructed-response items, and the intraclass correlation is reported for polytomously scored constructed-response items. A description of these measures is also included in *The NAEP 1996 Technical Report*.

5.6 DATA DELIVERY

The 1996 NAEP assessment data collection resulted in several classes of data files — student background, school, teacher, weights, SD/LEP student, student/teacher match, and student-response information. Student-response information included response data from all assessed students in 1996. Data resolution activities occurred prior to the submission of data files to ETS and Westat to resolve any irregularities that existed. This section details additional steps performed before creating the final data files to ensure capture of the most complete and accurate information.

An important quality-control component of the image-scoring system was the inclusion, with a student's response to one item, of an exact copy of the student edit record, including the student booklet ID number, with every image of a student's response to a constructed-response item. This information was used to identify the file within the image-scoring system. These edit files also remained in the main data files residing on the NCS mainframe computer. By attaching this information to a student's response, exact matching of scores assigned to constructed-response items and all other data for each individual student was guaranteed, because the booklet ID for each image was part of every image file. This ensured scores were applied to the correct student's record on the mainframe.

When all the responses for an individual item had been scored, the system automatically submitted all item scores assigned during scoring, along with their student edit records, to a queue to be transmitted to the mainframe. Project staff then initiated a system job to transmit all scoring data to be matched with the original student records on the mainframe. A custom edit program matched the edit records of the scoring files to those of the original edit records on the mainframe. As matches were confirmed, the scores were applied to those individual files. After completion of this stage, all data collected for an individual student was located in one single and complete record/file identified by the student edit record.

NCS processed the SD/LEP Student Questionnaires via OMR scanning. Edits performed on the questionnaires assured that responses to questions fell within the valid range for that question. SD/LEP questionnaires were then matched to a student record. SD/LEP questionnaires that were not matched to a student document were cross-referenced with the corresponding Administration Schedule, Roster of Questionnaire, and student data files to correct, if necessary, the information needed to result in a match.

In 1996, NCS continued to use ICR technology to capture percentage figures written by school personnel directly in boxes on the School Characteristics and Policies Questionnaires rather than requiring the school official to grid ovals in a matrix. The data were then verified by an edit operator.

The same processes that were followed in previous cycles were used in 1996 to achieve the best possible student/teacher match rate. The first step was to identify Teacher Questionnaires not returned to NCS for processing so as to exclude from the matching process the students of these teachers. Student identification numbers that were not matched to a Teacher Questionnaire were cross-referenced with the corresponding Administration Schedule and Roster of Teacher Questionnaires to verify (and change, if necessary) the teacher number, teacher period, and questionnaire number recorded on these control documents. The NAEP school numbers listed on the Roster of Questionnaires and Teacher Questionnaire were verified and corrected, if necessary. Once these changes were made, any duplicate teacher numbers existing within a school were, if possible, cross-referenced for resolution with the Rosters of Questionnaires. Because this information was located together on a single, central control document, the ability to match and resolve discrepant or missing fields was simplified.

After all data processing activities were completed, data cartridges and/or diskettes were created and shipped via overnight delivery to ETS and/or Westat, as appropriate. NCS maintains a duplicate archive file for security/backup purposes.

5.7 MISCELLANEOUS

5.7.1 Storage of Documents

After the batches of image-scanned documents had successfully passed the editing process, they were sent to the warehouse for storage. Due to the large number of rescore projects done with NAEP material, the documents were unspiraled and sequenced by grade and booklet type after all of the processing/scoring was completed. Unspiraled and sequenced booklets were then assigned a new inventory number by grade and booklet type and sent back to the warehouse for storage The storage locations of all documents were recorded on the inventory control system. Unused materials were sent to temporary storage to await completion of the entire assessment. Once the assessment was complete, NCS received authorization from ETS to salvage unused materials after determining that a sufficient quantity of each form type was retained permanently.

5.7.2 Quality Control Documents

ETS requires that a random sample of booklets and the corresponding scores/scoring sheets be pulled for an additional quality-control check that verifies the accuracy and completeness of the data. For image-scanned documents, a scoring sheet is not used, so ETS uses scores sent to them on a data tape to verify the accuracy of applied scores. All of these documents were selected prior to sending the booklets to storage. A random sample of all the questionnaires used in the 1996 NAEP assessment was also sent to ETS.

5.7.3 Alert Analysis

Table 5-10 identifies the different types of alerts to problems that were encountered in the processing of NAEP data. For the 1996 State Assessment, there was a total of 3,812 alerts.

Discrepancies were found in the receiving process that did not require an alert to be issued to Westat. They did require a great deal of effort by the opening staff to resolve in order to provide the most complete and accurate information. These are referred to as “info alerts.” These were categorized and codes were assigned to them. They are listed in the left-hand column of Table 5-10.

Even though receipt-control staff were well trained in the resolution of many situations, there were some problems that required resolution by NCS NAEP product line staff. These are referred to as “problem alerts.” The various types of problem alerts were also categorized and coded. They are listed in the right-hand column of Table 5-10. For any unusual situations, Westat was contacted to help with the resolution of the alert.

Table 5-10

Alerts for the 1996 National and State Assessments

Information Alerts	Problem Alerts
Code 52 not written on Administration Schedules	Change of Administration Codes-A/S or Booklets
The yes/no box not gridded on Rosters	Incorrect Rosters/Questionnaires
Session Number not on Administration Schedules	Administration Notes/Writing on Covers
Administration Codes not on A/S; but on booklets	Duplicate Student / Booklet Number/ Administration Schedule
Administration Codes not on booklets; but on A/S	All material not returned
Items returned for Westat	Affected Testing - Problem
Writing on booklet covers	Transcribed page(s) for student booklet(s)
Other	Processed as is
	Involves Inclusion Check List
	Other

A/S = Administration Schedules

Chapter 6

CREATION OF THE DATABASE, QUALITY CONTROL OF DATA ENTRY, AND CREATION OF THE DATABASE PRODUCTS[1]

John J. Ferris, Katharine E. Pashley, Patricia E. O'Reilly,
David S. Freund, and Alfred M. Rogers
Educational Testing Service

6.1 OVERVIEW

The data processing, scoring, and editing procedures described in Chapter 5 resulted in the generation of disk and tape files containing various data for students (assessed and excluded), teachers, and schools, along with SD/LEP (students with disabilities and students with limited English proficiency) information. The weighting procedures described in Chapter 7 resulted in the generation of data files that included the sampling weights required to make valid statistical inferences about the populations from which the 1996 State Assessment in science samples[2] were drawn. These files were merged into a comprehensive, integrated database. The creation of this database is described in Section 6.2.

Section 6.3 describes a central repository or master catalog of this information. The master catalog is accessible by all analysis and reporting programs and provides correct parameters for processing the data fields and consistent labeling for identifying the results of the analyses.

To evaluate the effectiveness of the quality control of the data entry process, the corresponding portion of the final integrated database was verified in detail against a sample of the original instruments received from the field. The results of this procedure are given in Section 6.4.

The integrated database was the source for the creation of the NAEP item information database and the NAEP secondary-use data files. These are described in Section 6.5.

6.2 MERGING FILES INTO THE STATE ASSESSMENT DATABASE

The data processing conducted by National Computer Systems (NCS) resulted in the transmittal to ETS of four data files for both fourth and eighth grade: one for the student background and item response data and one file for each of the three questionnaires (Science Teacher Questionnaire, School Characteristics and Policies Questionnaire, and SD/LEP

---

[1] John J. Ferris was responsible for the evaluation of the quality of the database and the date entry process; Katharine E. Pashley was responsible for database generation under the supervision of David S. Freund; Patricia E. O'Reilly and Alfred M. Rogers created the secondary-use data files.

[2] The 1996 State Assessment in science was conducted at grade 8 only, except for Department of Defense Education Activity (DoDEA) schools that were assessed at both grades 4 and 8. The sample design is described in detail in Chapter 3.

Questionnaire). The sampling weights, derived by Westat, Inc., comprised an additional seven files for each grade — three sets for assessed students, three sets for excluded students, and one for schools. (See Chapter 7 for a discussion of the sampling weights.) These 11 files at each grade were the foundation for the analysis of the 1996 State Assessment data. Before data analyses could be performed, these data files had to be integrated into a coherent and comprehensive database.

The 1996 State Assessment database for both fourth and eighth grade consisted of two files — student and school. Each record on the student file contained a student's responses to the particular assessment booklet the student was administered — Booklets 201 to 237 (in the case of excluded students, a booklet was assigned but the student response fields contain a special code indicating no response), and the information from the questionnaire that the student's science teacher completed. Additionally, for a student (assessed or excluded) who was identified as SD or LEP, the data from the SD/LEP Questionnaire is included. This questionnaire is filled out for all students identified as SD and/or LEP, both assessed and excluded. (See Chapter 2 for information regarding assessment instruments.) Also added to the student files were variables with school-level information supplied by Quality Education Data, Inc. (QED), including demographic information about schools such as race/ethnicity percentages. Because the teacher data is not a representative sample of teachers and as the focus of NAEP is to report student level results, the teacher response data was added to the student records. The school files were separate files that could be analyzed on their own and could also be linked to the student files through the unique school ID code.

The creation of the student data files for fourth and eighth grade began with the reorganization of the data files received from NCS. This involved two major tasks: 1) the files were restructured, eliminating unused (blank) areas to reduce the size of the files; and 2) in cases where students had chosen not to respond to an item, the missing responses were recoded as either "omitted" or "not reached," as discussed in Chapter 9. Next, the student response data were merged with the student weights files. The resulting file was then merged with the SD/LEP and teacher data. In all merging steps, the nine-digit booklet ID (the three-digit booklet number common to every booklet with the same block of items and a six-digit serial number unique to the booklet for a student) was used as the matching criterion.

The school file for each grade was created by merging the School Characteristics and Policies Questionnaire file with the file of school weights and school variables, supplied by Westat. The state and school codes were used as the matching criteria. Since some schools did not return a questionnaire, some of the records in the school file contained only school-identifying information and sampling weight information.

When the student and school files for each grade had been created, the database was ready for analysis. In addition, whenever new data values, such as composite background variables or scale scores, were derived, they were added to the appropriate database files using the same matching procedures described above.

For archival purposes and to provide data to the states, to researchers, and to policymakers, secondary-use data files and codebooks for each jurisdiction were generated from this database. The secondary-use data files, described in Section 6.5.2, contain all responses and response-related data from the assessment, including responses from the student booklets, Teacher Questionnaires, and School Characteristics and Policies Questionnaires, scale scores, sampling weights, and variables used to compute standard errors.

6.3 CREATING THE MASTER CATALOG

A critical part of any database is its processing control and descriptive information. Having a central repository of this information, which may be accessed by all analysis and reporting programs, will provide correct parameters for processing the data fields and consistent labeling for identifying the results of the analyses. The State Assessment master catalog file was designed and constructed to serve these purposes for the State Assessment database.

Each record of the master catalog contains the processing, labeling, classification, and location information for a data field in the State Assessment database. The control parameters are used by the access routines in the analysis programs to define the manner in which the data values are to be transformed and processed.

Each data field has a 50-character label in the master catalog describing the contents of the field and, where applicable, the source of the field. The data fields with discrete or categorical response values (e.g., multiple-choice items, professionally scored items, and most questionnaire items, but not weight fields) have additional label fields in the catalog containing 8- and 20-character labels for those response values. These shorter labels can be used for reporting purposes as a concise description of the responses for the items.

The classification area of the master catalog record contains distinct fields corresponding to predefined classification categories (e.g., science content and process areas) for the data fields. For a particular classification field, a nonblank value indicates the code of the subcategory within the classification categories for the data field. This classification area permits the grouping of identically classified items or data fields by performing a selection process on one or more classification fields in the master catalog.

The master catalog file was constructed concurrently with the collection and transcription of the State Assessment data so that it would be ready for use by analysis programs when the database was created. As new data fields were derived and added to the database, their corresponding descriptive and control information were entered into the master catalog. Machine-readable catalog files, created from the master catalog, are available as part of the secondary-use data files package for use in analyzing the data with programming languages other than SAS or SPSS (see Section 6.5.2.8). For SAS and SPSS users, files of control statements that create SAS or SPSS system files are provided (see Section 6.5.2.7).

6.4 QUALITY CONTROL EVALUATION

The purpose of the data entry quality control procedure is to gauge the overall accuracy of the process that transforms responses into machine-readable data. The procedure involves examining the actual responses made in a random sample of booklets and comparing them, mark by mark and character by character, with the responses recorded in the final database, which is used for analysis and reporting. Notwithstanding the marks made by the respondent, if the respondent's intention is unambiguous, and if the data entry system has failed to accurately capture the intended response, the erroneous data is considered a failure for purposes of this quality control evaluation.

The selection of booklets for this comparison took place at the point of first entry into the scanning process for data from the field. These selected quality control booklets were set aside in a predetermined proportion, using systematic random sampling, and then collected for subsequent close scrutiny. Selection proportions comparable to, or greater than, those used in previous assessments were used. The results of this process are discussed in detail below, and Table 6-2 contains detailed information about the sampling rates, numbers of booklets and data characters examined, and errors found.

6.4.1 Student Data

Thirty-seven assessment booklets, numbered 201 through 237, were administered to students as part of the State Assessment in science. Table 6-1 provides the numbers of each booklet in the database for each grade. Note that these numbers, and others reported below for various categories of data, may vary somewhat from other totals given in this report for a variety of reasons, having to do with the appropriateness of inclusion for different purposes. The variation in the numbers of student booklets is insignificant, according to a chi-square test, indicating very good control of the distribution process.

Student booklets were sampled in adequate numbers and the average rate of selection was 1 out of 385 or better, a selection rate comparable to that used in past assessments at both the state and national levels. The few errors found during this quality control examination did not cluster by booklet number, so there is no reason to believe that the variation in numbers of booklets selected had a significant effect on the estimates of overall error rate confidence limits reported below.

The quality control evaluation detected 34 errors in these student booklet samples, 4 at grade 4 (DoDEA samples only) and 30 at grade 8. Virtually all the errors involved either multiple responses that were not identified as such by the scanner or erasures that were recorded instead of ignored. To be considered a scanning error, the scanning process must have failed to correctly determine the respondent’s intent when it was plain to the human eye. While such a failure might seem to cast doubt on the scanning process, the final error rate determined from the quality control evaluation was reassuring. A very large volume of data was scanned with consistently usable results. An analysis of this evaluation based on the binomial theorem permits the inference of confidence limits indicated in the last column of Table 6-2; it is unlikely, for instance, that more than seventeen hundredths of a percent (.0017) of the data characters processed at grade 8 would differ from what a careful reader would have found in the student booklets. The seemingly much greater error rate for the grade 4 data is partially due to the small quality control sample taken; extrapolating from relatively few booklets leads to a more conservative estimate of the error rate.

6.4.2 Science Teacher Questionnaires

A total of 354 questionnaires from science teachers were associated with corresponding student data in the final database at grade 4 (DoDEA schools only), and 8,857 at grade 8. The grade 8 teacher questionnaires were sampled at the rate of about 1 in 100, the same rate used in the previous assessment. The grade 4 teacher questionnaires were sampled at a much greater rate, 1 in 59, because of the much smaller number of available questionnaires; only six were sampled, although the error rate estimated from these booklets was comparable to that from grade 8 (see Table 6-2). The 94 teacher questionnaires selected for quality control in this assessment contained a total of 57 errors in 31 different booklets. This is about two to three times the error rate found in recent assessments for this instrument.

While this error rate is not bad enough to render the teacher data unusable, there is some cause for concern in the fact that many of these errors were concentrated in a single item—the item requesting information on classroom size. This suggests the possible need for changing the presentation of this item in the questionnaire booklet. While such changes can cause unforeseen problems with response behavior and data reliability, it is being considered in this case because the design of this item appears to be unnecessarily complex and confusing to quite a few of the teachers. Secondary users of NAEP data are cautioned that although the classroom size was included as a conditioning variable in the analysis of these data, the errors found in the responses to this question make its use inadvisable. Other classroom data items also proved to be problematic for teachers, especially at Grade 8. The issue is currently undergoing study and alternatives are being considered.

6.4.3 School Characteristics and Policies Questionnaires

A total of 4,595 questionnaires were collected from school administrators and included in the database at grade 4, and 4,225 at grade 8. These questionnaires were sampled for quality control evaluation at the rate of about 1 in 75, resulting in the selection of 61 questionnaires at grade 4, and 54 at grade 8. The 17 errors that were found represent an error rate about the same as that for school questionnaires in past assessments, well below any reasonable threshold for alarm.

6.4.4 SD/LEP Student Questionnaires

A total of 396 SD/LEP questionnaires were scanned and included in the database at grade 4, and 12,670 at grade 8. Nearly half of these questionnaires represented students who were part of the cognitive assessment; the balance of the questionnaires came from students who were excluded. The overall selection rate was about 1 in 100, roughly double that used in earlier assessments for this type of questionnaire. A total of 131 questionnaires were selected across both grades. The resulting error rate indicated that the quality of this data was second only to the student data and certainly adequate for the purposes to which it was put.

The results of the evaluation of all questionnaire data, as well as the student data, are summarized in Table 6-2.

Table 6-1
Number of Science Booklets Scanned into Database and Selected for Quality Control Evaluation

Booklet Number	Number of Booklets in Database Grade 4	Number of Booklets in Database Grade 8	Number of Booklets Selected Grade 4	Number of Booklets Selected Grade 8
201	103	3,023	-	6
202	104	3,007	1	9
203	99	2,988	1	7
204	98	3,028	-	8
205	108	3,057	-	9
206	100	3,030	1	9
207	103	3,059	1	8
208	97	3,109	-	7
209	104	3,069	-	9
210	99	3,110	-	7
211	100	3,072	1	8
212	101	3,056	1	8
213	111	3,056	-	7
214	102	3,113	1	9
215	96	3,076	1	6
216	104	3,093	-	8
217	108	3,069	1	9
218	112	3,064	1	9
219	111	3,054	-	7
220	107	3,053	1	8
221	104	3,015	1	8
222	104	3,070	-	9
223	105	3,061	-	7
224	112	3,061	-	8
225	113	3,019	-	8
226	114	3,060	1	6
227	110	3,051	-	9
228	106	3,067	1	9
229	109	3,055	1	7
230	115	3,023	1	9
231	109	3,000	-	7
232	105	3,070	-	7
233	111	3,019	-	8
234	111	2,982	1	9
235	108	3,018	-	6
236	112	3,030	-	9
237	109	3,016	-	9
Total	**3,924**	**112,803**	**17**	**293**

Table 6-2
Summary of the Quality Control Evaluation of the Science Data

Subsample	Selection Rate	Different Booklets Selected	Number of Booklets Selected	Number of Characters of Data	Number of Errors	Observed Error Rate	Upper 99.8% Confidence Limit
GRADE 4:							
Student	1/231	37	17	1,493	4	.0027	.0093
Teacher	1/59	1	6	1,158	0	0	.0053
School[1]	1/75	1	61	10,858	11	.0010	.0023
SD/LEP	1/57	1	7	882	0	0	.0070
GRADE 8:							
Student	1/385	37	293	30,032	30	.0010	.0017
Teacher	1/101	1	88	15,840	57	.0036	.0052
School[1]	1/78	1	54	9,882	6	.0006	.0017
SD/LEP	1/102	1	124	15,624	17	.0011	.0021

[1]School figures are from the combined mathematics/science database since a school might be from one or both of these assessments.

6.5 NAEP DATABASE PRODUCTS

The NAEP database described to this point serves primarily to support analysis and reporting activities that are directly related to the NAEP cooperative agreement. This database has a singular structure and access methodology that is integrated with the NAEP analysis and reporting programs. One of the directives of the NAEP cooperative agreement is to provide secondary researchers with a nonproprietary version of the database that is portable to any computer system. In the event of transfer of NAEP to another client, the cooperative agreement further requires ETS to provide a full copy of the internal database in a format that may be installed on a different computer system.

In fulfillment of these requirements, ETS provides two sets of database products: the item information database and the secondary-use data files. The contents, format, and usage of these products are documented more extensively in the publications listed under the appropriate sections below.

6.5.1 The Item Information Database

The NAEP item information database contains all of the descriptive, processing, and usage information for each item or variable used for NAEP since 1970. The primary unit of this database is the item. Each NAEP item is associated with different levels of information, including usage across years and age cohorts, subject area classifications, response category descriptors, and locations of response data on secondary-use data files.

The item information database can be used for a variety of NAEP tasks: providing statistical information to aid in test construction, determining the usage of items across assessment years and ages for trend and cross-sectional analyses, providing text labels for analyses and reports, and organizing items by subject area classifications for scaling analysis.

6.5.2 The Secondary-Use Data Files

The secondary-use data files are designed to enable any researcher with an interest in NAEP to perform secondary analysis on the same data as those used for analysis at ETS. Supporting documentation accompanies the data files. The set of files for each sample (e.g., the North Dakota grade 8 assessed students) or instrument (e.g., the Florida grade 8 school data) includes: a file containing the data; a file of control statements that will generate an SPSS system file; a file of control statements that will generate a SAS system file; and a machine-readable catalog file. Each machine-readable catalog file (discussed in Section 6.5.2.8) contains sufficient control and descriptive information to aid those users without SAS or SPSS to set up and perform data analyses. The printed documentation consists of two volumes: a guide to the use of the data files, and a set of data file layouts and codebooks for each participating jurisdiction.

The remainder of this section summarizes the procedures used in generating the data files and related materials. More information about the contents and use of the data files is contained in the *NAEP 1996 State Assessment Program in Science Secondary-Use Data Files User Guide* (O'Reilly, Zelenak, Rogers, & Kline, 1998).

6.5.2.1 File Definition

There are essentially four samples for analysis in the 1996 State Assessment in science: the students (assessed and excluded), the schools in the State Assessment, and the students and the schools in a matched National Reporting Sample drawn from the national science assessment. The four samples are divided into separate files by participating jurisdiction (for the two State Assessment samples), resulting in a total of over 90 files; however, the same file formats, file linking conventions, and analysis considerations apply to each file within a given sample. For example, the analysis specification that links school and student data for California would apply identically to New York, Tennessee, or any other participating jurisdiction or group of jurisdictions.

Every data file for each participating jurisdiction requires its own data codebook, detailing the frequencies of data values within that jurisdiction for the given sample. The file layouts, SPSS and SAS syntax, and machine-readable catalog files, however, need only be generated for each of the four samples, since the individual jurisdiction data files for each 1996 State Assessment sample are identical in format and data code definition.

6.5.2.2 Definition of the Variables

Prior to the 1990 assessment, information which could potentially be used to identify students or schools was not included on the secondary-use files. When these public-use data files were replaced by the current restricted-use data files, the restraint on confidential data was lifted. This change simplified the variable definition process, as it permitted the transfer of *all* variables from the database to the secondary-use files.

The initial step in this process was the generation of a LABELS file of descriptors of the variables for each data file to be created. Each record in a LABELS file contains, for a single data field, the variable name, a short description of the variable, and processing control information to be used by later steps in the process of generating the secondary-use data files. ETS staff could edit this file for deletion of variables, modification of control parameters, or reordering of the variables within the file. The LABELS file is an intermediate file only; it is not distributed with the secondary-use data files.

The next program in the processing stream, GENLYT, produced a printed layout for each data file from the information in its corresponding LABELS file. These layouts are reviewed for the ordering of the variables. The variables on all data files are grouped and arranged in the following order: identification information, weights, derived variables, scale scores (where applicable), and item response data. On the student data files, these fields are followed by the teacher response data and the SD/LEP student questionnaire data, where applicable. The identification information is taken from the front covers of the instruments. The weight data include sample descriptors, selection probabilities, and replicate weights for the estimation of sampling error. The derived data include sample descriptions from other sources and variables that are derived from the item response data for use in analysis or reporting. Item response data consist of responses to questionnaire items; for assessed students, these data include responses to cognitive items, as well.

In the assessed student data files for each participating jurisdiction of the State Assessment in science and for the National Reporting Sample, the item response data within each block were left in their order of presentation. The blocks, however, were arranged according to the following scheme: common background, subject-related background, the cognitive blocks in ascending numerical order, and student motivation. The responses to cognitive blocks that were not present in a given booklet were left blank, signifying a condition of 'missing by design.'

In order to process and analyze the spiral sample data effectively, the user must also be able to determine, from a given booklet record, which blocks of item response data were present and their relative order in the instrument. The user obtains this information from a set of control variables, one for each block, which indicate not only the presence or absence of the block but its order in the instrument. These control variables created by ETS are included with the derived variables.

6.5.2.3 Data Definition

To enable the data files to be processed on any computer system using any procedural or programming language, it was desirable that the data be expressed in numeric format. This was possible, but not without the adoption of certain conventions for expressing the data values numerically.

During creation of the NAEP database, the responses to all multiple-choice items (both cognitive multiple-choice items and those in the questionnaires) were processed and stored in the database using the letter codes printed in the instruments. This scheme afforded the advantage of saving storage space for items with 10 or more response options, but at the expense of translating these codes into their numeric equivalents for analysis purposes. The response data fields for most of these items would require a simple alphabetic-to-numeric conversion. However, the data fields for items with 10 or more response choices would require "expansion" before the conversion, since the numeric value would require two column positions. One of the processing control parameters on the LABELS file indicates whether or not the data field is to be expanded before conversion and output to the secondary-use data files.

The ETS database contained special codes to indicate certain response conditions: "I don't know" responses, multiple responses, omitted responses, not-reached responses, and unresolvable responses, which included out-of-range responses and responses that were missing due to errors in printing or processing. The scoring guides for the science constructed-response items included additional special codes for ratings of erased or crossed out and for ratings of illegible, "I don't know," off task, or nonratable by the scorers. All of these codes had to be reexpressed in a consistent numeric format.

The following convention was adopted and used in the designation of these codes: The "I don't know" and nonratable response codes (including off-task and illegible responses) were always converted to 7; the omitted response codes were converted to 8; the not-reached response codes were converted to 9; the multiple response codes were converted to 0; and the erased and crossed out response codes were converted to 5. The out-of-range and missing responses were coded as blank fields, corresponding to the 'missing by design' designation.

This coding scheme created conflicts for those multiple-choice items that had seven or more valid response options as well as the "I don't know" response, and also for those constructed-response items whose scoring guide had five or more categories. These data fields were also expanded to accommodate the valid response values and the special codes. In these cases, the special codes were 'extended' to fill the output data field: the "I don't know" and nonratable codes were extended from 7 to 77, omitted response codes from 8 to 88, etc.

Each numeric variable on the secondary-use files was classified as either continuous or discrete. These classifications are related to machine-level characteristics, rather than to the precise mathematical meaning of these terms. The discrete variables include those items for which each numeric value corresponds to a response category. The continuous variables include the weights, scale scores, the identification information codes, and questionnaire item responses for which counts or percentages were requested. The designation of "discrete" includes those derived variables to which numeric classification categories have been assigned. The constructed-response items were treated as a special subset of the discrete variables and were assigned to a separate category to facilitate their identification in the documentation.

6.5.2.4 Data File Catalogs

The LABELS file contains sufficient descriptive information for generating a brief layout of the data file. However, to generate a complete codebook document, substantially more information about the data is required. The CATALOG file provides most of this information.

The CATALOG file is created by the GENCAT program from the LABELS file and the 1996 master catalog file, as described in Section 6.3. Each record on the LABELS file generates a CATALOG record by first retrieving the master catalog record corresponding to the field name. The master catalog record contains usage, classification, and response code information, along with positional information from the LABELS file: field sequence number, output column position, and field width. Like the LABELS file, the CATALOG file is an intermediate file and is not included with the secondary-use data files.

The information for the response codes, also referred to as "foils," consists of the valid data values for the discrete numeric fields and a 20-character descriptive label for each valid data value. (Readers who are familiar with standard usage of the term "foil" in testing and measurement will notice that it has an expanded meaning in this discussion of the secondary-use data files.) The GENCAT program uses additional control information from the LABELS file to determine if extra foils should be generated and saved with each CATALOG record. The first flag controls generation of the "I don't know" or nonratable foil; the second flag regulates omitted or not-reached foil generation; and the third flag denotes the possibility of multiple responses for that field and sets up an appropriate foil. All of these control parameters, including the expansion flag, may be altered in the LABELS file by use of a text editor, in order to control the generation of data or descriptive information for any given field.

The LABELS file supplies control information for many of the subsequent secondary-use data processing steps. The CATALOG file provides detailed information for those and other steps.

6.5.2.5 Data File Layouts

The data file layouts, as mentioned above, were the first user product to be generated in the secondary-use data files process. The generation program, GENLYT, used a LABELS file, described in Section 6.5.2.2, and a CATALOG file as input and produced a printable file. The LAYOUT file is basically a formatted listing of the LABELS file; it documents the layout and contents of the data files. The layouts are part of the printed documentation; the secondary-use data file package includes not only the printed layouts, but also the electronic files from which they were printed.

Each line of the LAYOUT file contains the following information for a single data field: sequence number, field name, output column position, field width, number of decimal places, data type, value range, key or correct response value, and a short description of the field. The sequence number of each field is implied from its order on the LABELS file. The field name is an 8-character label for the field that is used consistently by all secondary-use data file materials to refer to that field on that file. The output column position is the relative location of the beginning of that field on each record for that file, using bytes or characters as the unit of measure. The field width indicates the number of columns used in representing the data values for a field. If the field contains continuous numeric data, the value under the number of decimal places entry indicates how many places to shift the decimal point before processing data values.

The data type category uses five codes to designate the nature of the data in the field: Continuous numeric data are coded “C”; discrete numeric data are coded “D”; constructed-response item data are coded “OS” if the item was dichotomized for scaling and “OE” if it was scaled under a polytomous response model. Additionally, the discrete numeric fields that include “I don’t know” response codes are coded “DI.” If the field type is discrete numeric, the value range is listed as the minimum and maximum permitted values separated by a hyphen to indicate range. If the field is a response to a multiple-choice item, the correct option value, or key, is printed; if the field is an assigned score for a constructed-response item that was scaled as a dichotomous item using cutpoint scoring, the range of correct scores is printed. Each variable is further identified by its 50-character descriptive label.

6.5.2.6 Data Codebooks

The data codebooks form the bulk of the printed documentation of the secondary-use data files; they contain complete descriptive information for each data field. Most of this information originates from the CATALOG file; the remaining data comes from the COUNTS file and the IRT parameters file, described below. The secondary-use data file package includes the electronic files from which the codebooks were printed, in addition to the printed codebooks.

Each data field receives at least one line of descriptive information in the codebook. If the data type is continuous numeric, no more information is given. If the variable is discrete numeric, the codebook lists the foil codes, foil labels, and frequencies of each value in the data file. Additionally, if the field represents an item used in IRT scaling, the codebook lists the final parameters estimated by the scaling program. (See Chapters 8 and 9 for information about scaling.)

Certain blocks of cognitive items in the 1996 assessment that are to be used again in later assessments for trend comparisons have been designated as nonreleased. In order to maintain confidentiality of nonreleased multiple-choice items, generic foil labels have been substituted for the foils (i.e., the response category descriptions) for these items in the data codebooks and the secondary-use files.

The frequency counts are not available on the CATALOG file, but must be generated from the data. The GENFREQ program creates the COUNTS file using the field name to locate the variable in the database, and the foil values to validate the range of data values for each field. This program also serves as a check on the completeness of the foils in the CATALOG file, as it flags any data values not represented by a foil value and label.

The IRT parameter file is linked to the CATALOG file through the field name. Printing of the IRT parameters is governed by a control flag in the classification section of the CATALOG record. If an item has been scaled, and, thus, used in deriving the scale scores, the IRT parameters are listed to the right of the foil values and labels, and the score value for each response code is printed to the immediate right of the corresponding frequency.

The LAYOUT and CODEBOOK files are written by their respective generation programs to print-image disk data files. Draft copies are printed and distributed for review before the production copy is generated. The production copy is printed on an IBM printer that uses laser-imaging technology to produce high-quality, reproducible documentation.

6.5.2.7 Control Statement Files for Statistical Packages

An additional requirement of the NAEP cooperative agreement is to provide, for each secondary-use data file, a file of SAS control statements that will convert the secondary-use data file into a system data file for use with the SAS statistical system. Also required is a file of SPSS control statements that will produce a system data file for the SPSS statistical system. Two separate programs, GENSAS and GENSPX, generate these control statement files using the CATALOG file as input.

The control statement files create a SAS or SPSS system data file that corresponds to an entire NAEP secondary-use data file. NAEPEX, the NAEP data extraction software described in Section 6.5.2.9, can be used to produce control statement files that create a SAS or SPSS system data file corresponding to a user-defined subset of the NAEP secondary-use data files. Also described in that section are the NAEP analysis modules, currently available for use with SPSS® for Windows™.

Each of the control statement files contains separate sections for variable definition, variable labeling, missing value declaration, value labeling, and creation of scored variables from the cognitive items. The variable definition section describes the locations of the fields, by name, in the file, and, if applicable, the number of decimal places or type of data. The variable label identifies each field with its 50-character descriptive label. The missing value section identifies values of those variables that are to be treated as missing and excluded from analyses. The value labels correspond to the foils in the CATALOG file. The code values and their descriptors are listed for each discrete numeric variable. The scoring section is provided to permit secondary users to generate item score variables in addition to the item response variables.

Each of the control statement generation programs combines three steps into one complex procedure. As each CATALOG file record is read, it is broken into several component records according to the information to be used in each of the resultant sections. These component records are tagged with the field sequence number and a section sequence code. They are then sorted by section code and sequence number. Finally, the reorganized information is output in a structured format dictated by the syntax of the processing language.

ETS tests the control statement files by using them to generate system data files from the secondary-use data files. The control statement files are distributed in the secondary-use data files package to permit users with access to SAS and/or SPSS to create their own system data files.

6.5.2.8 Machine-Readable Catalog Files

For those NAEP data users who have neither SAS nor SPSS capabilities, yet require processing control information in a computer-readable format, the distribution files also contain machine-readable catalog files. Each machine-readable catalog record contains processing control information, IRT parameters, and foil codes and labels.

6.5.2.9 Secondary-Use Data Files on CD-ROM

The complete set of secondary-use data files described above are available on CD-ROM as part of the NAEP Data on Disk product suite. This medium can be used by researchers and policy makers operating in a personal computing environment.

The NAEP Data on Disk product suite includes two additional components which facilitate the analysis of NAEP secondary-use data. The PC-based NAEP data extraction software, NAEPEX, enables users to create customized extracts from the NAEP secondary-use data files and to generate SAS or SPSS control statements for preparing analyses or generating customized system files. Both Windows 3.1 and DOS versions of NAEPEX are available. The NAEP analysis modules, which currently run under SPSS® for Windows™, use output files from the extraction software to perform analyses that incorporate statistical procedures appropriate for the NAEP design.

Summarized NAEP data in tabular format (the NAEP data almanacs described in Chapter 10) are also available on CD-ROM. This product, which is distinct from the secondary-use data files, includes the NAEP almanac viewer, a program that allows users to locate and display data of

Chapter 7

WEIGHTING PROCEDURES AND VARIANCE ESTIMATION[1]

Penny James and John Burke
Westat, Inc.

7.1 OVERVIEW

Following the collection of assessment and background data from and about assessed and excluded students, the processes of deriving sampling weights and associated sets of replicate weights were carried out. The sampling weights are needed to make valid inferences from the student samples to the respective populations from which they were drawn. Replicate weights are used in the estimation of sampling variance, through the procedure known as *jackknife repeated replication*.

Each student was assigned a weight to be used for making inferences about the state's students. This weight is known as the *full-sample* or *overall* sample weight. The full-sample weight contains three components. First, a base weight is established that is the inverse of the overall probability of selection of the sampled student. The base weight incorporates the probability of selecting a school and the student within a school. This weight is then adjusted for two sources of nonparticipation — school level and student level. These weighting adjustments seek to reduce the potential for bias from such nonparticipation by increasing the weights of students from schools similar to those schools not participating, and increasing the weights of students similar to those students from within participating schools who did not attend the assessment session (or makeup session) as scheduled. The details of how these weighting steps were implemented are given in Sections 7.2 and 7.3.

Section 7.4 addresses the effectiveness of the adjustments made to the weights using the procedures described in Section 7.3. The section examines characteristics of nonresponding schools and students, and investigates the extent to which nonrespondents differ from respondents in ways not accounted for in the weight adjustment procedures. Section 7.5 considers the distributions of the final student weights in each jurisdiction, and whether there were outliers that called for further adjustment.

In addition to the full-sample weights, a set of replicate weights was provided for each student. These replicate weights are used in calculating the sampling errors of estimates obtained from the data, using the jackknife repeated replication method. Full details of the method of using these replicate weights to estimate sampling errors are contained in the *Technical Report of the NAEP 1994 Trial State Assessment Program in Reading* (Mazzeo, Allen, & Kline, 1995) and in earlier state technical reports. Section 7.6 of this report describes how the sets of replicate weights were generated for the 1996 State Assessment data. The methods of deriving these weights were aimed at reflecting the features of the sample design appropriately in each

[1] In addition to his responsibility in the sampling activities, John Burke was responsible for directing all weighting and variance estimation procedures. Penny James contributed by carrying out most of these procedures.

jurisdiction, so that when the *jackknife variance estimation* procedure is implemented, approximately unbiased estimates of sampling variance result.

As detailed in Chapter 4, two different sets of inclusion rules indicated by the sample type field were used in the 1996 State Assessment program. To enable ETS to analyze these subsets separately, the student weights for each subset were raked in order to force agreement with the totals estimated using both subsets combined. This *raking* process is detailed in Section 7.7.

7.2 CALCULATION OF BASE WEIGHTS

7.2.1 Calculation of School Base Weights

The base weight assigned to a school w_i^{sch} was the reciprocal of the probability of selection of that school. For the eighth-grade samples and fourth-grade Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) and Department of Defense Dependents Schools (DoDDS), the school base weight depended on the subject of assessment because some schools were so small that students were tested in only one subject. For "new" schools selected using the supplemental new school sampling procedures (see Chapter 3), the school base weight reflected the combined probability of selection of the district, and school within district.

In each jurisdiction, all schools included in the sample with certainty were assigned school base weights of unity. Schools sampled with certainty were sometimes selected more than once in the systematic sampling process. For example, a school that was selected twice was allocated twice the usual number of students for the assessments, or two sessions; a school that was selected three times was allocated three times the usual number of students for the assessments, or three sessions. All schools at grade 8 and DDESS and DoDDS schools at grade 4 that had fewer than 20 students were assigned one subject (See Chapter 3). For these schools, the base weight included a factor of 2. Additional details about the weighting process are given in the sections below.

7.2.2 Weighting New Schools

New public schools were identified and sampled through a two-stage sampling process, involving the selection of districts, and then of new schools within selected districts. This process is described in Chapter 3. There were two distinct processes used depending upon the size of the district.

Within each jurisdiction, public school districts were partitioned into "small" districts, which are those having at most three schools on the aggregate frame and no more than one fourth-, one eighth-, and one twelfth-grade school. The remainder of the districts were denoted as "medium" or "large" districts. For the larger districts (those having multiple schools in at least one of grades 4, 8, and 12), a sample of districts was selected in each jurisdiction. Districts in the sample were asked to identify schools having grade 4 or grade 8 that were not included on the school frame. A sample of these newly identified schools was then selected. The base weight for these schools reflected both the probability that the district was selected for this updating process, and that the school was included in the NAEP sample, having been identified as new by the district. If the school was in grade 8 or grade 4 DDESS and DoDDS schools, but was only large enough to do one subject, the base weight included a factor of 2 as described in Section 7.2.1.

There were no schools identified in small districts (see Tables 3-8 and 3-9).

7.2.3 Treatment of Substitute and Double-Session Substitute Schools

Schools that replaced a refusing school (i.e., substitute schools) were assigned the weight of the refusing school. Thus the substitute school was treated as if it were the original school that it replaced, for purposes of obtaining school base weights. Schools conducting extra sessions that served as substitutes for a refusing school (i.e., double-session substitutes) in effect had two school weights. The students in the school who were assigned to the original session were given the school base weight of the participating school, while those students assigned to the extra session(s) were assigned the school base weight of the refusing school. The base weight was adjusted by a factor of 2 if the grade 8 or DDESS or DoDDS school was only large enough to do one subject.

7.2.4 Calculation of Student Base Weights

Within the sampled schools, eligible students were sampled for assessment using the procedures described in Chapter 3. The within-school probability of selection for science therefore depended on the number of grade-eligible students in the school and the number of students selected for the assessment (usually 30). The within-school weights for the substitute schools were further adjusted to compensate for differences in the sizes of the substitute and the originally sampled (replaced) schools. In the case of the fourth grade DDESS and DoDDS schools and all eighth-grade schools, the within-school weight also reflected the fact that a small school could have been selected for one subject but not the other. Thus, in general, the within-school student weight for the *j*th student in school *i* was equal to:

$$W_{ij}^{within} = \frac{N_i}{n_i} K_{1i} \; K_{2i}$$

where

N_i = the number of grade-eligible students enrolled in the school, as reported in the sampling worksheets; and

n_i = the number of students selected for the given subject.

The factors K_{1i} and K_{2i} in the formula for the within-school student weight generally apply to only a few schools in each jurisdiction. The factor K_{1i} adjusts the count of grade-eligible students in a substitute school to be consistent with the corresponding count of the originally sampled (replaced) school. Specifically, for substitute schools,

$$K_{1i} = \frac{E_i}{E_i^s}$$

with

E_i = the grade enrollment of the originally sampled (replaced) school; and

E_i^s = the grade enrollment of the substitute school.

For nonsubstitute schools, $K_{1i} = 1$.

The factor K_{2i}, that was applied to schools determined to be “year-round” schools, is defined as

$$K_{2i} = \frac{1}{1 - p_{off}}$$

where p_{off} is the percentage of students enrolled in the school who were not scheduled to attend school at the time of assessment. For schools that are not year-round schools (the great majority), $K_{2i} = 1$.

The overall student base weight for a student j selected for the science assessment in school i was obtained by multiplying the school base weight by the within-school student weight and therefore was computed as:

$$W_{ij}^{base} = W_i^{sch} \; W_{ij}^{within} .$$

7.3 ADJUSTMENTS FOR NONRESPONSE

As mentioned earlier, the base weight for a student was adjusted by two factors: one to adjust for nonparticipating schools for which no substitute participated, and one to adjust for students who were invited to the assessment but did not appear in the scheduled sessions (original or makeup).

7.3.1 Defining Initial School-Level Nonresponse Adjustment Classes

School-level nonresponse adjustment classes were created separately for public and nonpublic schools within each jurisdiction. For each set these classes were defined as a function of their sampling strata, as follows.

Public Schools. For each jurisdiction, except Guam, the initial school nonresponse adjustment classes were formed by crossclassifying the level of urbanization and minority status (see Chapter 3 for definitions of these characteristics). Where there were no minority strata within a particular level of urbanization, a categorized version of median household income was used. For this purpose within each level of urbanization, public schools were sorted by the median household income, and then divided into three groups of about equal size, representing low, middle, and high income areas. In Guam, where there was no information on minority status or median household income, grade enrollment was used.

DDESS and DoDDS Schools. For the jurisdictions comprised of DDESS and DoDDS schools, the initial nonresponse adjustment classes were defined by other geographic variables. For the DDESS schools, the classes were defined by military installation and grouped by nearby jurisdictions. For DoDDS, schools were grouped by the regions of Europe or the Pacific where the military installation was located.

Nonpublic Schools. For each jurisdiction (excluding District of Columbia and Guam nonpublic schools), initial nonresponse adjustment classes were formed by crossclassifying school type (Catholic and non-Catholic) and metropolitan status (metro/nonmetro). For District of Columbia nonpublic schools, these classes were defined by crossclassifying school type and two levels of estimated grade enrollment (25 or fewer students, versus 26 or more students). For Guam, initial nonresponse classes for nonpublic schools were defined by school type only. The District of Columbia is entirely metropolitan, and Guam is entirely nonmetropolitan, so alternatives were needed for these two jurisdictions.

7.3.2 Constructing the Final Nonresponse Adjustment Classes

The objective in forming the nonresponse adjustment classes is to create as many classes as possible that are internally as homogeneous as possible, but such that the resulting nonresponse adjustment factors are not subject to large random variation. Consequently, all initial nonresponse adjustment classes deemed unstable were collapsed with suitable neighboring classes so that: (1) the combined class contained at least six sessions, and (2) the resulting nonresponse adjustment factor did not exceed 1.35 (in a few cases a factor in excess of 1.35 was permitted). These limits had been used for the 1994 Trial State Assessment. One change was implemented for the 1996 State Assessment. When 100 percent of the public schools in a jurisdiction responded, no action was taken for a public-school adjustment class that contained fewer than six sessions. The same approach was used for nonpublic schools where 100 percent of them participated. Although clearly there is no adjustment for school nonresponse in these cases, this change in procedure could have an effect on the final definition of the student nonresponse adjustment classes (Section 7.3.4).

Public Schools. For these schools, inadequate nonresponse adjustment classes were reinforced by collapsing adjacent levels of minority status (or median household income level if minority information was missing). In doing so, metropolitan and nonmetropolitan schools were not mixed. All DDESS and DoDDS schools cooperated, so no collapsing of schools was necessary.

Nonpublic Schools. For nonpublic schools, excluding schools in District of Columbia and Guam, inadequate classes were reinforced by collapsing adjacent levels of metropolitan-area status. Catholic and non-Catholic schools were kept apart to the extent possible, particularly when the only requirement to combine such schools was as a means of reducing the adjustment factors below 1.35. For schools in the District of Columbia, inadequate classes were collapsed over similar values of estimated grade enrollment. Catholic and non-Catholic schools were kept apart to the extent possible. For nonpublic schools in Guam, Catholic and non-Catholic schools were collapsed together in order to form a stable nonresponse adjustment class.

7.3.3 School Nonresponse Adjustment Factors

The school-level nonresponse adjustment factor for the *i*th school in the *h*th class was computed as:

$$F_h^{(1)} = \frac{\sum_{i \in C_h} W_{hi}^{sch} E_{hi}}{\sum_{i \in C_h} W_{hi}^{sch} E_{hi} \delta_{hi}}$$

where

C_h = the subset of school records in class h;

W_{hi}^{sch} = the base weight of the ith school in class h;

E_{hi} = the grade enrollment for the ith school in class h;

$$\delta_{hi} = \begin{cases} 1 & \text{if the } i\text{th school in adjustment class } h \\ & \text{participated in the assessment; and} \\ 0 & \text{otherwise} \end{cases}$$

Both the numerator and denominator of the nonresponse adjustment factor contained only schools that were determined to have eligible students enrolled.

In the calculation of the above nonresponse adjustment factors, a school was said to have participated if:

- It was selected for the sample from the frame or from the lists of new schools provided by participating school districts, and student assessment data were obtained from the school; or

- The school participated as a substitute school and student assessment data were obtained (so that the substitute participated in place of the originally selected school).

The nonresponse-adjusted weight for the *i*th school in class *h* was computed as:

$$W_{hi}^{adj} = F_h^{(1)} W_{hi}^{sch} \quad .$$

7.3.4 Student Nonresponse Adjustment Classes

The initial student nonresponse adjustment classes for assessed students were formed based on several variables. The first of these was public/nonpublic strata. Public/nonpublic strata were then crossclassified by a variable created from combining SD/LEP status and the sample type for the student. SD denotes students with disabilities, while LEP denotes students classified as having limited English proficiency. Within these categories, the initial student nonresponse adjustment classifications were defined separately depending on the SD/LEP status of a student.

If a student was SD or LEP, then the class was formed by urbanicity crossclassified by student age. Age was used to classify students into two groups (for grade 4, those born in September 1985 or earlier and those born in October 1985 or later, and for grade 8, those born in September 1981 or earlier and those born in October 1981 or later). If a student was neither SD nor LEP, then the initial nonresponse adjustment class was formed by urbanicity crossclassified by student age (as defined above), by the quality control monitoring status (see Chapter 3), then finally by minority status as collapsed for the school nonresponse. For the DDESS and DoDDS schools, the nonresponse adjustment classes for SD and LEP students were student age crossclassified by the geographic variable as defined for the school nonresponse adjustment classes.

Following creation of the initial student nonresponse adjustment classifications, all unstable classes were identified for possible collapsing with other classes. A class was considered to be unstable when either of the following conditions was true for the given class:

- Number of responding eligible students was fewer than 20; or
- Nonresponse adjustment factor exceeded 1.5.

All classes deemed unstable in the previous step were collapsed with other classes using the following rules:

- Do not collapse across public and nonpublic;
- Do not collapse across SD/LEP and non-SD/non-LEP;
- If within cells defined by the crossclassification of public/nonpublic and SD-LEP/nonSD-nonLEP status, and sample type within the SD/LEP categories, all of the adjustments are one, no adjustments are made; and
- Collapse across the last variable of the nonresponse adjustment cell only (i.e., collapse across geography for SD/LEP students in Department of Defense Education Activity (DoDEA) schools).

More collapsing was necessary only if the resulting classes had fewer than 15 responding eligible students. Collapsing then continued within the successive variables until the class size was no longer deficient or until a “set” boundary that could not be crossed was reached.

7.3.5 Student Nonresponse Adjustments

As described above, the student-level nonresponse adjustments for the assessed students were made within classes defined by the SD/LEP status, sample type, final school-level nonresponse adjustment classes, monitoring status of the school, and age group of the students. Subsequently, in each jurisdiction, the final student weight for the *j*th student of the *i*th school in class *k* was then computed as:

$$W_{kij}^{final} = W_i^{adj} \times W_{ij}^{within} \times F_k$$

where

W_i^{adj} = the nonresponse-adjusted school weight for school *i*;

W_{ij}^{within} = the within-school weight for the *j*th student in school *i*;

and

$$F_k = \frac{\sum_j W_{ij}}{\sum_j W_{ij}\,\delta_{kj}} \; .$$

In the above formulation, the summation included all students, *j*, in the *k*th final (collapsed) nonresponse class. The indicator variable δ_{kj} had a value of 1 when the *j*th student in adjustment class *k* participated in the assessment; otherwise, δ_{kj} = 0.

For excluded students, no nonresponse adjustment procedures were applied because excluded students were not required to complete an assessment. In effect, all excluded students were considered respondents. Weights are provided for excluded students so as to estimate the size of this group and its population characteristics. Tables 7-1 through 7-3 summarize the unweighted and final weighted counts of assessed and excluded students in public and nonpublic schools for grade 8 and for DDESS and DoDDS schools at both grades 4 and 8.

Table 7-1
Unweighted and Final Weighted Counts of Assessed and Excluded Students
Grade 4 DoDEA Schools

	Assessed		**Excluded**		**Assessed and Excluded**	
Jurisdiction	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**
DoDEA/DDESS	1,293	2,771	88	178	1,381	2,949
DoDEA/DoDDS	2,631	6,478	123	290	2,754	6,768
Total	**3,924**	**9,248**	**211**	**468**	**4,135**	**9,717**

Table 7-2

Unweighted and Final Weighted Counts of Assessed and Excluded Students by Jurisdiction
Grade 8 Public Schools

	Assessed		Excluded		Assessed and Excluded	
Jurisdiction	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**
Alabama	2,267	51,331	204	4,186	2,471	55,517
Alaska	1,590	8,152	84	362	1,674	8,514
Arizona	2,218	52,997	179	3,872	2,397	56,869
Arkansas	1,894	33,432	172	2,780	2,066	36,211
California	2,452	343,278	280	34,503	2,732	377,781
Colorado	2,607	46,519	183	2,962	2,790	49,482
Connecticut	2,585	32,974	269	3,113	2,854	36,087
Delaware	1,943	7,990	105	354	2,048	8,344
District of Columbia	1,737	4,137	190	385	1,927	4,522
Florida	2,473	144,418	294	15,973	2,767	160,391
Georgia	2,507	84,236	165	5,117	2,672	89,353
Guam	953	2,178	62	121	1,015	2,299
Hawaii	2,232	12,115	152	795	2,384	12,910
Indiana	2,370	71,973	166	4,520	2,536	76,493
Iowa	2,260	35,237	131	1,923	2,391	37,160
Kentucky	2,536	45,555	113	1,876	2,649	47,431
Louisiana	2,682	53,738	192	3,492	2,874	57,229
Maine	2,344	15,727	169	1,033	2,513	16,760
Maryland	2,179	54,196	150	3,053	2,329	57,249
Massachusetts	2,418	59,071	192	4,132	2,610	63,204
Michigan	2,236	110,734	140	6,035	2,376	116,770
Minnesota	2,459	60,431	113	2,573	2,572	63,004
Mississippi	2,518	38,537	193	2,637	2,711	41,174
Missouri	2,472	58,170	175	3,803	2,647	61,973
Montana	2,092	13,107	88	439	2,180	13,545
Nebraska	2,824	22,857	120	917	2,944	23,773
Nevada	1,006	18,649	75	1,416	1,081	20,065
New Hampshire	1,790	14,504	114	788	1,904	15,292
New Jersey	1,625	75,291	184	7,933	1,809	83,224
New Mexico	2,494	21,111	288	2,170	2,782	23,281
New York	1,966	179,460	189	15,523	2,155	194,983
North Carolina	2,703	85,512	143	4,124	2,846	89,636
North Dakota	2,590	9,235	66	229	2,656	9,464
Oregon	2,380	36,724	143	1,800	2,523	38,524
Rhode Island	2,190	9,920	183	748	2,373	10,668
South Carolina	2,215	46,811	166	3,122	2,381	49,932
Tennessee	2,367	59,172	125	2,914	2,492	62,086
Texas	2,396	246,946	233	21,477	2,629	268,423
Utah	2,764	36,935	158	1,839	2,922	38,774
Vermont	2,022	7,448	129	450	2,151	7,898
Virginia	2,624	77,866	219	5,753	2,843	83,619
Washington	2,586	69,590	118	2,861	2,704	72,451
West Virginia	2,649	22,665	233	1,817	2,882	24,482
Wisconsin	2,186	63,912	188	4,967	2,374	68,879
Wyoming	2,709	7,684	104	293	2,813	7,978
Total	**102,110**	**2,552,525**	**7,339**	**191,182**	**109,449**	**2,743,706**

Table 7-3

Unweighted and Final Weighted Counts of Assessed and Excluded Students by Jurisdiction
Grade 8 Nonpublic Schools

	Assessed		Excluded		Assessed and Excluded	
Jurisdiction	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**
Alabama	144	4,541	0	0	144	4,541
Arkansas	89	1,435	0	0	89	1,435
California	208	31,789	0	0	208	31,789
Connecticut	270	4,627	2	32	272	4,659
Delaware	315	1,620	1	5	316	1,625
District of Columbia	266	1,078	0	0	266	1,078
DoDEA/DDESS	620	1,338	36	72	656	1,410
DoDEA/DoDDS	2,268	4,849	77	154	2,345	5,003
Georgia	238	6,080	5	30	243	6,110
Guam	199	535	0	0	199	535
Iowa	247	3,284	0	0	247	3,284
Kentucky	263	4,836	0	0	263	4,836
Louisiana	428	11,012	2	54	430	11,067
Maryland	322	8,999	2	31	324	9,030
Massachusetts	340	10,011	3	112	343	10,122
Michigan	332	14,603	4	187	336	14,790
Minnesota	249	5,258	3	45	252	5,304
Missouri	366	7,816	1	20	367	7,836
Montana	154	691	2	7	156	697
Nebraska	337	2,760	1	6	338	2,766
Nevada	133	983	2	12	135	994
New Hampshire	188	1,160	2	10	190	1,170
New Jersey	294	13,497	2	87	296	13,584
New Mexico	230	1,960	0	0	230	1,960
New York	516	34,374	3	161	519	34,535
North Dakota	162	675	1	3	163	679
Oregon	54	3,200	0	0	54	3,200
Rhode Island	340	1,662	3	14	343	1,676
South Carolina	138	3,348	0	0	138	3,348
Texas	130	13,587	0	0	130	13,587
Utah	96	1,017	1	11	97	1,028
Vermont	115	365	2	7	117	372
Washington	215	5,332	0	0	215	5,332
Wisconsin	380	11,193	2	57	382	11,250
Wyoming	47	158	0	0	47	158
Total	**10,693**	**219,674**	**157**	**1,116**	**10,850**	**220,790**

7.4 CHARACTERISTICS OF NONRESPONDING SCHOOLS AND STUDENTS

In the previous section, procedures were described for adjusting the survey weights so as to reduce the potential bias of nonparticipation of sampled schools and students. To the extent that a nonresponding school or student is different from those respondents in the same nonresponse adjustment class, potential for nonresponse bias remains.

In this section we examine the potential for remaining nonresponse bias in two related ways. First, we examine the weighted distributions, within each jurisdiction at grade 8, of certain characteristics of schools and science students, both for the full sample and for respondents only. This analysis is of necessity limited to those characteristics that are known for both respondents and nonrespondents, and hence cannot directly address the question of nonresponse bias. The approach taken does reflect the reduction in bias obtained through the use of nonresponse weighting adjustments. As such, it is more appropriate than a simple comparison of the characteristics of nonrespondents with those of respondents for each jurisdiction.

The second approach involves modeling the probability that a school is a nonrespondent, as a function of the nonresponse adjustment class within which the school is located, together with other school characteristics. This has been achieved using linear logistic regression models, with school response status as the dependent variable. By testing to see if the school characteristics add any predictive ability to the model, over and above using the membership of the nonresponse adjustment class to make this prediction, we can obtain some insight into the remaining potential for nonresponse bias. If these factors are substantially marginally predictive, there is a danger that significant nonresponse bias remains. These models have been developed for public schools in each of 11 jurisdictions at grade 8 having public school participation (after substitution) of below 90 percent (with a participation rate prior to substitution in excess of 70%).

7.4.1 Weighted Distributions of Schools Before and After School Nonresponse

Table 7-4 shows the mean values of certain school characteristics for public schools, both before and after nonresponse. The means are weighted appropriately to reflect whether nonresponse adjustments have been applied (i.e., to responding schools only) or not (to the full set of in-scope schools). The variables for which means are presented are the percentage of students in the school who are Black, the percentage who are Hispanic, the median household income (1989) of the ZIP code area where the school is located, and the type of location. All variables were obtained from the sample frame, and so from Quality Education Data, Inc. (QED), described in Chapter 3, with the exception of the type of location. This variable was derived for each sampled school using U.S. Bureau of Census data. The type of location variable has seven possible levels, which are defined in Chapter 3. Although this variable is not interval-scaled, the mean value does give an indication of the degree of urbanization of the population represented by the school sample (lower values for type of location indicate a greater degree of urbanization).

Table 7-4

Weighted Mean Values Derived from Sampled Public Schools - Grade 8

Jurisdiction	Weighted Participation Rate After Substitution (%)	Weighted Mean Values Derived from Full Sample				Weighted Mean Values Derived from Responding Sample, with Substitutes and School Nonresponse Adjustment			
		Percent Black	Percent Hispanic	Median Income	Type of Location	Percent Black	Percent Hispanic	Median Income	Type of Location
Alabama	89.57	35.76	0.20	$23,855	4.90	35.60	0.23	$23,669	4.90
Alaska	93.25	4.81	2.03	$36,324	4.77	4.77	2.03	$36,355	4.77
Arizona	87.37	4.14	24.95	$30,729	3.14	3.83	26.00	$30,322	3.16
Arkansas	70.51	22.65	0.34	$22,472	5.43	22.25	0.36	$22,561	5.45
California	94.08	7.73	33.15	$36,539	3.29	7.77	32.73	$36,418	3.26
Colorado	100.00	5.94	16.24	$32,422	3.68	5.94	16.24	$32,422	3.68
Connecticut	100.00	11.71	8.77	$45,934	3.84	11.71	8.77	$45,934	3.84
Delaware	100.00	27.50	3.02	$35,376	5.04	27.50	3.02	$35,376	5.04
District of Columbia	100.00	89.10	5.65	$29,035	1.00	89.10	5.65	$29,035	1.00
Florida	100.00	23.49	14.67	$28,821	3.33	23.49	14.67	$28,821	3.33
Georgia	99.00	28.60	1.14	$30,466	4.39	28.53	1.14	$30,568	4.39
Guam	100.00	3.19	0.53	-----	7.00	3.19	0.53	-----	7.00
Hawaii	100.00	2.40	5.45	$35,001	4.20	2.40	5.45	$35,001	4.20
Indiana	89.63	12.07	2.05	$28,365	4.46	12.27	2.10	$28,513	4.45
Iowa	82.85	3.04	1.13	$27,735	5.00	3.16	1.05	$27,827	5.01
Kentucky	92.28	8.69	0.09	$23,663	5.32	8.91	0.15	$23,807	5.33
Louisiana	100.00	42.14	0.96	$23,505	4.52	42.14	0.96	$23,505	4.52
Maine	90.63	0.14	0.50	$29,242	5.69	0.15	0.53	$29,290	5.69
Maryland	85.59	32.29	2.49	$41,211	3.60	32.58	2.34	$41,372	3.60
Massachusetts	92.38	8.05	7.67	$41,890	4.03	8.27	7.81	$42,416	4.03
Michigan	86.78	16.16	1.84	$33,498	4.28	16.66	1.96	$33,464	4.29
Minnesota	88.32	3.83	1.32	$32,800	4.89	3.79	1.23	$32,931	4.89
Mississippi	95.37	49.62	0.07	$20,965	5.64	49.36	0.07	$20,995	5.64
Missouri	95.95	14.65	0.66	$28,400	4.76	14.63	0.68	$28,507	4.77
Montana	75.85	0.28	1.08	$24,892	5.51	0.30	1.05	$24,479	5.52
Nebraska	99.88	5.59	2.85	$28,547	5.00	5.59	2.86	$28,552	5.00
Nevada	38.05	9.74	13.13	$32,297	3.82	11.10	14.48	$31,594	3.71
New Hampshire	68.40	0.89	0.84	$39,134	5.28	0.92	0.65	$39,012	5.34
New Jersey	63.53	17.00	11.68	$43,776	3.71	17.90	14.47	$42,288	3.66
New Mexico	100.00	2.30	45.05	$24,187	4.73	2.30	45.05	$24,187	4.73
New York	77.67	19.63	13.99	$35,173	3.32	18.53	13.33	$35,665	3.35
North Carolina	100.00	30.67	0.77	$28,071	4.89	30.67	0.77	$28,071	4.89
North Dakota	93.32	0.63	0.66	$27,620	5.18	0.66	0.67	$27,429	5.11
Oregon	91.87	2.14	4.44	$29,789	3.91	2.35	4.56	$29,755	3.91
Rhode Island	89.96	6.05	6.63	$32,559	3.54	6.15	7.10	$32,234	3.52
South Carolina	86.70	41.49	0.39	$26,878	4.98	41.60	0.38	$27,022	4.98
Tennessee	91.73	20.99	0.19	$25,615	4.41	20.00	0.21	$26,041	4.41
Texas	96.11	12.88	33.24	$28,374	3.53	13.03	33.01	$28,195	3.50
Utah	100.00	0.60	3.67	$32,265	4.34	0.60	3.67	$32,265	4.34
Vermont	74.90	0.54	0.18	$31,784	6.13	0.55	0.16	$31,334	6.06
Virginia	100.00	21.10	1.69	$38,695	4.30	21.10	1.69	$38,695	4.30
Washington	95.27	4.05	5.72	$34,722	4.08	4.06	5.82	$34,650	4.08
West Virginia	100.00	4.02	0.06	$22,622	5.48	4.02	0.06	$22,622	5.48
Wisconsin	78.30	7.75	1.89	$32,196	4.63	8.34	2.05	$31,769	4.64
Wyoming	100.00	0.75	5.95	$31,374	5.38	0.75	5.95	$31,374	5.38

Two sets of means are presented for these four variables. The first set shows the weighted mean derived from the full sample of in-scope schools selected for science; that is, respondents and nonrespondents (for which there was no participating substitute). The weight for each sampled school is the product of the school base weight and the grade enrollment. This weight therefore represents the number of students in the state represented by the selected school. The second set of means is derived from responding schools only, after school substitution. In this case the weight for each school is the product of the nonresponse-adjusted school weight and the grade enrollment of the original school, and therefore indicates the number of students in the jurisdiction represented by the responding school.

Table 7-5 shows some of these same statistics for all schools combined, for those jurisdictions where both the public-school participation rate prior to substitution, and the nonpublic-school participation rate prior to substitution, exceeded 70 percent. These are the jurisdictions for which assessment results have been published for both public and nonpublic schools combined. Data on minority enrollment were not available for nonpublic schools, and so are not included in Table 7-5.

The differences between these sets of means give an indication of the potential for nonresponse bias that has been introduced by nonresponding schools for which there was no participating substitute. For example, in New York the mean percentage Black enrollment, estimated from the original sample of public schools, is 19.63 percent (Table 7-4). The estimate from the responding schools is 18.53 percent. Thus there may be a slight bias in the results for New York because these two means differ. Note, however, that throughout these four tables the differences in the two sets of mean values are generally very slight, at least in absolute terms, suggesting that it is unlikely that substantial bias has been introduced by schools that did not participate and for which no substitute participated. Of course in a number of states (as indicated) there was no nonresponse at the school level, so that these sets of means are identical. Even in those jurisdictions where school nonresponse was relatively high (such as Arkansas and Vermont), the absolute differences in means are slight. Occasionally the relative difference is large (the “Percent Black” in Nevada public schools, for example), but these are for small population subgroups, and thus are very unlikely to have a large impact on results for the jurisdiction as a whole. Also, this is for a jurisdiction with low participation (38.05%) for which data would not be published because of concern about nonresponse bias.

7.4.2 Characteristics of Schools Related to Response

In an effort to evaluate the possibility that substantial bias remains as a result of school nonparticipation, following the use of nonresponse adjustments, a series of analyses were conducted on the response status for public schools. This analysis was restricted to those jurisdictions with a participation rate below 90 percent (after substitution), because these are the jurisdictions where the potential for nonresponse bias is likely to be the greatest. Those jurisdictions with an initial public-school response rate below 70 percent were not included, since NAEP does not report results for these jurisdictions because of concern about nonresponse bias. Information about this can be found in Chapter 10. Nonpublic schools were omitted from these analyses because of the small sample sizes involved, which means that it is difficult to assess whether a potential for bias exists.

Table 7-5
Weighted Mean Values Derived from All Sampled Schools for Jurisdictions Achieving Minimal Required Public and Nonpublic School Participation, Grade 8

	Weighted Participation Rate After Substitution (%)		Weighted Mean Values Derived from Full Sample		Weighted Mean Values Derived from Responding Sample, with Substitutes and School Nonresponse Adjustment	
Jurisdiction	Public	Nonpublic	Median Income	Type of Location	Median Income	Type of Location
Arkansas	70.51	73.53	$22,922	5.37	$23,070	5.39
California	94.08	79.63	$36,947	3.28	$36,923	3.25
Georgia	99.00	87.69	$30,657	4.30	$30,814	4.29
Guam	100.00	78.78	-----	7.00	-----	7.00
Iowa	82.85	93.72	$27,790	4.88	$27,841	4.88
Kentucky	92.28	81.80	$24,303	5.16	$24,462	5.15
Louisiana	100.00	74.60	$24,230	4.24	$24,096	4.29
Massachusetts	92.38	77.48	$41,864	3.91	$41,984	3.87
Michigan	86.78	86.75	$33,600	4.18	$33,475	4.17
Minnesota	88.32	83.94	$32,944	4.76	$32,955	4.80
Missouri	95.95	100.00	$29,675	4.59	$29,767	4.60
Montana	75.85	97.03	$24,869	5.47	$24,479	5.47
Nebraska	99.88	84.18	$28,532	4.85	$28,480	4.89
New Mexico	100.00	95.18	$24,111	4.66	$24,044	4.67
New York	77.67	87.31	$34,878	3.09	$35,062	3.09
North Dakota	93.32	78.16	$27,618	5.15	$27,424	5.09
Texas	96.11	78.83	$28,473	3.47	$28,314	3.44
Vermont	74.90	80.42	$31,836	6.07	$31,430	5.99
Washington	95.27	85.56	$34,796	3.97	$34,788	3.97
Wyoming	100.00	92.34	$31,330	5.37	$31,333	5.37

The 11 states investigated were the following (with the public school participation rate shown in parentheses): Arizona (87%), Arkansas (71%), Iowa (83%), Maryland (86%), Michigan (87%), Minnesota (88%), Montana (76%), New York (78%), South Carolina (87%), Vermont (75%), and Wisconsin (78%). The approach used was to develop a logistic regression model within each jurisdiction, to predict the probability of participation as a function of the nonresponse adjustment classes and other school characteristics. The aim was to determine whether the response rates are significantly related to school characteristics, after accounting for the effect of the nonresponse class. Thus "dummy" variables were created to indicate nonresponse class membership.

If there are *k* nonresponse classes within a jurisdiction, let

X_{ij} = 1 if the school j is classified in nonresponse class i
= 0 otherwise, for i = 1,...,(k-1)

Within each jurisdiction a logistic model was fitted to the data on public-school participation. In the model, the indicator variables for nonresponse class were included, and also additional variables available for participating and nonparticipating schools alike. These variables were the percentage of Black students (Y_1), the percentage of Hispanic students (Y_2), the estimated grade 8 enrollment size of the school (Y_3), and the median household income of the ZIP code area in which the school was located (Y_4).

The model fitted in each jurisdiction was the following:

$$L_j = A + \Sigma B_i X_{ij} + \Sigma C_i Y_{ij} \ .$$

Let P_j denote the probability that school j is a participant, and let L_j denote the logit of $P_{j.}$ That is,

$$L_j = ln(P_j \ / \ (1 - P_j \ )).$$

Note that this model cannot be estimated if there are nonresponse classes in which all schools participated (so that no adjustments for nonresponse were made for schools in such a class). Even though this analysis was restricted to those jurisdictions with relatively poor response, this occurred in a number of instances. When this happened, those (responding) schools in such classes were dropped from the analyses. Table 7-6 shows the proportion of the state public-school student population that is represented in the sample by schools from classes with less than 100 percent response. Thus for Iowa and Vermont, there was some nonresponse within every adjustment class, whereas for the other nine jurisdictions some portion of the population is not represented because schools were dropped from classes with no nonresponse.

The table shows that only two of the models that contained all of the variables were significant. These were the models for Arizona and Arkansas, both of which had p-values less than .001. For Arizona, the only individual variable for this model that was significant was median income. This variable was not used in forming nonresponse.

For Arkansas, the only individual variable for this model that was slightly significant was the dummy variable corresponding to nonresponse class 2, which indicates for this state that the nonresponse classes significantly explain the variation in the response rates. Nonresponse class 2 was comprised of schools with high income in midsize central cities and urban fringes of large or midsize cities.

Table 7-6
Results of Logistic Regression Analysis of School Nonresponse - Grade 8

Jurisdiction	School Participation Rate (%)	Percent of Population Covered by Model	Model with All Variables: Degrees of Freedom	Model with All Variables: Significance	Model with All Variables: Significant Variables	Test: Y_{ij}'s = 0: Degrees of Freedom	Test: Y_{ij}'s = 0: Significance
Arizona	87.37	83.67	12	**p<.001**	median p=.001	4	**p<.001**
Arkansas	70.51	92.58	7	**p<.001**	nonresponse cell 2 p=.016	4	p=.791
Iowa	82.85	100.00	11	p=.807	none	4	p=.540
Maryland	85.59	84.46	7	p=.141	none	4	p=.425
Michigan	86.78	72.01	7	p=.367	median p=.017	4	p=.115
Minnesota	88.32	70.31	9	p=.996	none	4	p=.909
Montana	75.85	79.70	7	p=.725	none	4	p=.966
New York	77.67	69.10	7	p=.555	none	4	p=.284
South Carolina	86.70	76.31	9	p=.659	none	4	p=.873
Vermont	74.90	100.00	5	p=.467	none	4	p=.690
Wisconsin	78.30	91.07	9	p=.646	est. enrollment p=.018	4	p=.184

Median household income is somewhat significant for Michigan. This variable was used in forming nonresponse adjustment classes in Michigan in large/small towns and rural areas. Minority enrollment was used in other areas (see Appendix F). This significance does not translate into the results of Table 7-4, since the median household income for the full sample is $33,498, which is very close to the value of $33,464 for the respondents.

Another variable that is significant is the estimated grade 8 enrollment. This variable is significant for Wisconsin. For public schools, this variable was not used in forming nonresponse adjustment classes in Wisconsin (it was used only in Guam). This variable is not shown in Table 7-4. However, the sign of the coefficient for this variable in the logistic model indicates that smaller schools are somewhat under-represented in Michigan.

To determine if the variables other than the nonresponse adjustment class variables added explanatory power to the model, all variables except the nonresponse adjustment class variables were tested collectively to see if the estimates of the parameters were equal to zero. This evaluates whether, taken as a group, the Y variables are significantly related to the response probability, after accounting for nonresponse class. The results are shown in the last columns of Tables 7-6. Only one of the tests was significant (for Arizona). This suggests that for this state, the variables did add to the model after accounting for the nonresponse class, considering that the significant variable for this model was median household income.

However, all of the results for the other models suggest that the variables did not add to the model after accounting for the nonresponse adjustment classes, even though on occasion an individual variable was significant. These results hold for Arkansas, where the full model was significant. For Arkansas, the significant variable was the nonresponse class 2, which also indicates for this state that the additional variables did not add significantly to the model.

These results indicate that on occasion there are differences between the original samples of schools and those that participated, that are not fully removed by the process of creating nonresponse adjustments. Although these effects are not dramatic, they are sometimes statistically significant, and may be reflected in noticeable differences in population characteristics estimated from the respondents, compared to those obtained for the full sample. However, the evidence presented here does not permit valid speculation about the likely size or even direction of the bias in science achievement results in the states where these sample differences are noticeable.

7.4.3 Weighted Distributions of Students Before and After Student Absenteeism

Tables 7-7 and 7-8 show, for the public schools in grade 8 in each jurisdiction and the DDESS and DoDDS jurisdictions in both grades 4 and 8, the weighted sampled percentages of students by gender (male) and race/ethnicity (White, not Hispanic; Black, not Hispanic; Hispanic), SD/LEP status for the full sample of students (after student exclusion), and for the assessed sample. The mean student age in months is also presented on each basis. Table 7-9 shows these results for all students, public and nonpublic, in those jurisdictions at grade 8 having adequate school response rates to permit reporting of combined results for public and nonpublic students.

Table 7-7

Weighted Student Percentages Derived From Sampled DoDEA Schools - Grade 4

Jurisdiction	Weighted Student Participation (%)	Weighted Estimates Derived from Full Sample: Percent Male	Percent White	Percent Black	Percent Hispanic	Percent SD	Percent LEP	Mean Age (Months)	Weighted Estimates Derived from Assessed Sample, with Student Nonresponse Adjustment: Percent Male	Percent White	Percent Black	Percent Hispanic	Percent SD	Percent LEP	Mean Age (Months)
DoDEA/DDESS	95.59	51.71	46.59	26.88	19.72	5.26	0.29	128.25	51.61	45.80	27.32	20.06	5.24	0.31	128.32
DoDEA/DoDDS	94.37	50.07	44.64	17.45	19.62	4.34	1.16	127.72	50.03	44.11	17.40	20.32	4.29	1.21	127.76

Table 7-8

Weighted Student Percentages Derived From Sampled Public and DoDEA Schools - Grade 8

		Weighted Estimates Derived from Full Sample							Weighted Estimates Derived from Assessed Sample, with Student Nonresponse Adjustment						
Jurisdiction	**Weighted Student Participation (%)**	**Percent Male**	**Percent White**	**Percent Black**	**Percent Hispanic**	**Percent SD**	**Percent LEP**	**Mean Age (Months)**	**Percent Male**	**Percent White**	**Percent Black**	**Percent Hispanic**	**Percent SD**	**Percent LEP**	**Mean Age (Months)**
Alabama	92.70	49.57	60.96	33.32	3.19	5.70	0.05	179.72	48.98	61.77	32.27	3.25	5.80	0.05	179.66
Alaska	81.59	51.05	66.20	4.41	6.34	8.63	3.09	177.24	51.14	65.75	3.98	7.41	8.50	3.26	177.17
Arizona	89.75	49.67	57.90	4.04	29.74	3.98	4.83	178.63	49.83	57.32	3.78	30.58	3.91	4.95	178.67
Arkansas	92.48	49.85	73.81	19.89	3.80	3.91	0.49	179.53	50.19	73.12	20.15	4.10	4.03	0.49	179.43
California	91.61	49.43	37.10	8.08	40.16	3.79	9.62	175.93	49.42	38.30	7.28	38.90	3.74	9.25	175.87
Colorado	91.41	49.86	69.71	5.04	20.12	5.43	1.04	177.66	49.95	69.69	4.90	20.24	5.42	1.05	177.59
Connecticut	92.61	49.69	74.48	10.01	11.39	6.57	0.74	176.63	49.35	74.78	9.93	11.05	6.57	0.73	176.66
Delaware	89.24	51.59	63.75	26.43	6.90	6.25	0.21	177.08	50.58	64.04	25.59	7.15	6.26	0.21	176.96
District of Columbia	84.77	49.14	3.01	83.67	10.01	3.08	0.88	178.16	49.08	2.92	82.56	11.02	2.98	0.99	177.65
DoDEA/DDESS	94.82	52.26	48.61	21.50	23.97	4.84	0.00	176.27	52.64	47.12	22.29	24.52	4.84	0.00	176.08
DoDEA/DoDDS	93.36	49.48	45.90	18.05	16.03	3.67	1.19	176.38	48.94	45.12	18.45	16.57	3.71	1.15	176.38
Florida	90.02	52.91	55.36	20.41	21.14	7.54	1.93	179.16	52.78	54.96	20.23	21.60	7.40	2.07	179.07
Georgia	91.63	49.70	56.26	35.89	4.67	4.32	0.28	179.88	50.15	56.04	35.78	4.91	4.35	0.24	179.82
Guam	89.64	50.82	7.70	2.26	16.99	2.67	1.46	176.37	50.20	7.87	2.54	19.20	2.67	1.46	176.35
Hawaii	89.79	52.03	17.56	2.81	20.30	5.10	1.98	174.95	51.65	17.18	3.11	22.34	5.17	1.93	174.78
Indiana	92.17	50.75	80.83	11.26	5.16	4.94	0.07	179.71	50.20	80.73	10.99	5.45	4.96	0.08	179.63
Iowa	93.51	49.84	90.59	3.58	3.27	8.99	0.17	178.77	49.75	90.97	3.17	3.21	8.59	0.17	178.70
Kentucky	94.44	50.46	86.30	8.56	3.12	5.59	0.21	179.53	50.48	85.75	8.69	3.36	5.53	0.22	179.38
Louisiana	90.18	50.04	55.24	37.12	5.10	4.63	0.40	179.81	49.84	54.84	37.03	5.47	4.58	0.45	179.65
Maine	92.01	47.83	92.97	1.29	2.82	6.93	0.49	179.32	48.66	92.52	1.25	3.04	6.96	0.46	179.33
Maryland	88.96	51.23	55.82	32.96	5.82	6.73	0.59	175.36	50.90	55.77	32.12	6.39	6.67	0.65	175.19
Massachusetts	91.49	51.90	80.28	6.84	7.95	10.05	1.13	177.39	51.73	80.48	6.60	7.80	9.92	1.26	177.32
Michigan	89.79	50.26	72.93	17.67	4.39	3.93	0.66	178.00	50.08	75.71	14.87	4.55	3.94	0.63	177.79
Minnesota	91.54	50.24	84.72	5.07	3.94	6.97	0.29	178.58	50.04	85.25	4.37	4.10	7.00	0.30	178.56
Mississippi	91.63	51.23	49.79	43.87	4.97	4.48	0.16	181.91	50.30	49.57	43.44	5.61	4.45	0.10	181.62
Missouri	91.77	50.77	77.82	14.39	4.59	6.78	0.28	179.23	50.63	78.31	13.37	4.84	6.81	0.31	179.24
Montana	91.70	49.63	83.28	0.70	4.89	6.28	0.11	179.57	49.25	83.37	0.66	5.17	5.91	0.13	179.56
Nebraska	91.97	49.68	85.56	5.06	6.37	7.07	0.37	178.42	50.10	85.50	5.06	6.53	7.25	0.38	178.40
Nevada	92.38	48.43	60.16	6.67	24.03	3.89	2.17	177.39	48.94	59.80	6.52	24.27	4.24	2.29	177.35
New Hampshire	89.92	52.45	90.68	1.31	4.88	9.14	0.10	179.04	53.32	90.04	1.51	5.21	9.23	0.12	179.00
New Jersey	93.38	46.74	60.55	14.80	17.31	5.45	0.78	176.51	46.79	60.22	14.40	17.86	5.35	0.85	176.42
New Mexico	90.40	50.20	37.77	2.32	49.98	8.61	3.29	178.09	50.07	38.01	2.40	49.92	8.61	3.31	178.06
New York	89.94	50.26	55.02	20.29	17.73	5.71	2.78	176.05	50.47	58.71	17.07	16.55	5.15	2.56	176.16
North Carolina	91.34	49.93	64.67	27.06	4.01	5.04	0.23	178.02	49.98	64.54	26.58	4.42	5.02	0.26	177.88
North Dakota	93.79	52.00	91.24	1.22	3.33	6.51	0.27	178.97	52.12	91.31	1.19	3.68	6.60	0.31	179.00
Oregon	89.22	49.42	81.11	2.31	8.01	7.04	1.33	177.87	49.35	81.31	2.11	8.24	6.97	1.22	177.96
Rhode Island	88.75	50.43	78.15	5.22	10.86	8.26	2.11	176.93	49.64	77.06	5.33	11.62	8.10	2.27	176.82
South Carolina	90.35	49.45	51.52	40.51	5.48	4.35	0.11	179.10	49.56	50.88	40.38	6.11	4.29	0.12	179.03
Tennessee	91.24	51.91	76.97	17.91	2.97	7.76	0.08	179.35	51.96	77.01	17.40	3.23	7.74	0.10	179.23
Texas	92.24	50.71	47.73	12.69	36.25	6.21	3.89	179.60	50.50	48.51	12.24	35.82	6.05	3.71	179.42
Utah	90.45	48.44	87.00	0.65	8.20	4.07	0.30	176.89	48.18	86.69	0.67	8.43	4.04	0.30	176.91
Vermont	93.08	49.35	90.43	1.14	3.75	8.79	0.34	177.10	49.25	89.73	1.06	4.15	8.75	0.39	177.20
Virginia	90.30	50.93	63.99	25.28	4.74	4.44	1.69	177.19	50.80	64.12	24.86	4.98	4.38	1.69	177.17
Washington	90.09	51.06	74.75	4.18	9.70	5.54	1.44	177.88	51.09	74.55	4.11	9.87	5.52	1.48	177.84
West Virginia	92.63	51.09	90.35	3.93	2.93	4.96	0.07	178.57	51.32	90.02	3.79	3.23	4.96	0.08	178.43
Wisconsin	90.45	49.98	83.20	7.39	5.11	3.95	0.49	178.56	49.51	83.50	6.66	5.48	3.96	0.47	178.53
Wyoming	93.13	51.36	83.59	0.58	10.87	6.05	0.65	178.92	51.47	83.42	0.63	10.95	6.04	0.66	178.89

Table 7-9

Weighted Student Percentages Derived From All Schools Sampled

Public and Nonpublic Schools - Grade 8

	Weighted Student Participation (%)		Weighted Estimates Derived from Full Sample							Weighted Estimates Derived from Assessed Sample, with Student Nonresponse Adjustment						
Jurisdiction	Public	Nonpublic	Percent Male	Percent White	Percent Black	Percent Hispanic	Percent SD	Percent LEP	Mean Age (Months)	Percent Male	Percent White	Percent Black	Percent Hispanic	Percent SD	Percent LEP	Mean Age (Months)
Arkansas	92.48	99.03	50.35	74.26	19.18	3.99	3.79	0.47	179.47	50.66	73.58	19.44	4.28	3.91	0.47	179.37
California	91.61	96.02	49.46	38.78	7.86	38.64	3.56	8.70	175.85	49.46	39.72	7.14	37.61	3.54	8.47	175.80
Georgia	91.63	96.33	49.72	58.28	33.93	4.71	4.06	0.26	179.74	50.18	58.09	33.80	4.95	4.10	0.23	179.68
Guam	89.64	94.28	50.71	8.13	2.09	15.98	2.14	1.27	175.94	50.21	8.27	2.35	17.90	2.14	1.27	175.92
Iowa	93.51	95.91	50.30	90.79	3.33	3.29	8.20	0.16	178.72	50.11	91.08	2.98	3.25	7.92	0.15	178.65
Kentucky	94.44	96.68	50.20	86.60	8.11	3.17	5.15	0.19	179.25	50.29	86.05	8.27	3.40	5.11	0.20	179.14
Louisiana	90.18	95.57	48.63	59.35	32.71	5.01	4.39	0.33	179.01	48.46	58.95	32.66	5.32	4.34	0.37	178.87
Massachusetts	91.49	94.05	52.67	81.26	6.29	7.55	8.98	1.08	177.20	52.33	81.42	6.10	7.40	8.79	1.19	177.14
Michigan	89.79	96.74	49.80	73.61	17.03	4.84	3.64	0.57	177.92	49.80	76.13	14.50	4.87	3.66	0.55	177.74
Minnesota	91.54	94.35	50.24	85.35	4.76	3.85	6.44	0.27	178.49	50.18	85.82	4.13	4.01	6.47	0.27	178.47
Missouri	91.77	95.17	51.32	79.24	13.08	4.46	6.25	0.25	179.12	50.94	79.66	12.18	4.68	6.28	0.27	179.13
Montana	91.70	92.86	49.79	82.72	0.84	5.14	5.96	0.11	179.53	49.34	82.94	0.80	5.34	5.62	0.45	179.56
Nebraska	91.97	96.50	49.91	86.24	4.63	6.22	6.56	0.34	178.46	50.25	86.23	4.59	6.37	6.67	0.34	178.43
New Mexico	90.40	95.33	50.31	37.64	2.22	48.60	8.07	3.01	178.11	50.18	37.88	2.30	48.60	8.07	3.03	178.07
New York	89.94	96.92	50.39	52.55	22.89	17.92	4.83	2.29	175.83	50.60	55.92	19.86	16.83	4.43	2.15	175.93
North Dakota	93.79	93.43	52.73	90.94	1.14	3.25	6.38	0.67	178.94	52.66	90.81	1.10	3.57	6.45	0.63	178.99
Texas	92.24	97.72	51.03	48.13	12.15	36.30	5.99	3.68	179.63	50.80	48.87	11.76	35.85	5.85	3.52	179.45
Vermont	93.08	91.02	49.44	90.62	1.12	3.68	8.38	0.33	177.06	49.37	89.93	1.05	4.06	8.35	0.37	177.16
Washington	90.09	94.70	50.73	75.15	4.28	9.20	5.14	1.37	177.86	50.81	74.92	4.21	9.38	5.13	1.41	177.82
Wyoming	93.13	93.94	51.62	82.69	0.59	10.88	6.03	0.64	178.95	51.70	82.56	0.65	10.97	6.02	0.65	178.93

The weight used for the full sample is the adjusted student base weight, defined in Section 7.2.4. The weight for the assessed students is the final student weight, defined in Section 7.3.5. The difference between the estimates of the population subgroups is an estimate of the bias in estimating the size of the subgroup, resulting from student absenteeism.

Care must be taken in interpreting these results, however. First, note that there is generally very little difference in the proportions estimated from the full sample and those estimated from the assessed students. While this is encouraging, it does not eliminate the possibility that bias exists, either within the state as a whole, or for results for gender and race/ethnicity subgroups, or for other subgroups. Second, on the other hand, where differences do exist they cannot be used to indicate the likely magnitude or direction of the bias with any reliability. For example, in Table 7-8, for New York the percentages of Black and Hispanic students in the full sample are respectively 20.29 and 17.73 percent. For assessed students, these percentages are 17.07 for Black students and 16.55 for Hispanic students. While these differences raise the possibility that some bias exists, it is not appropriate to speculate on the magnitude of this bias by considering the assessment results for Black and Hispanic students, in comparison to other students in the state. This is because the underrepresented Black and Hispanic students may not be typical of students that were included in the sample, and similarly those students within the same racial/ethnic groups who are disproportionately overrepresented may not be typical either. The reason is that not all students within the same race/ethnicity group receive the same student nonresponse adjustment.

One other feature to note is that, for assessed students, information as to the student's gender and race/ethnicity is provided by the student, while for absent students this information is provided by the school. Evidence from past NAEP assessments (see, for example, Rust & Johnson, 1992) indicates that there can be substantial discrepancies between those two sources, especially with regard to classifying grade 4 students as Hispanic.

7.5 VARIATION IN WEIGHTS

After computation of full-sample weights, an analysis was conducted on the distribution of the final student weights in each jurisdiction and for DDESS and DoDDS schools at each grade. The analysis was intended to 1) check that the various weight components had been derived properly in each jurisdiction, and 2) examine the impact of the variability of the sample weights on the precision of the sample estimates, both for the jurisdiction as a whole and for major subgroups within the jurisdiction.

The analysis was conducted by looking at the distribution of the final student weights for the assessed students in each jurisdiction separately by public and nonpublic schools. Two key aspects of the distribution were considered in each case: the coefficient of variation (equivalently, the relative variance) of the weight distribution; and the presence of outliers-that is, cases whose weights were several standard deviations away from the median weight.

It was important to examine the coefficient of variation of the weights because a large coefficient of variation reduces the effective size of the sample. Assuming that the variables of interest for individual students are uncorrelated with the weights of the students, the sampling variance of an estimated average or aggregate is approximately $(1+\left[\frac{C}{100}\right]^2)$ times as great as the corresponding sampling variance based on a self-weighting sample of the same size, where C is the coefficient of variation of the weights expressed as a percent. Outliers, or cases with extreme weights, were examined because the presence of such an outlier was an indication of the possibility that an error was made in the weighting procedure, and because it was likely that a few extreme cases would contribute substantially to the size of the coefficient of variation.

In most jurisdictions, the coefficients of variation were 35 percent or less, both for the whole sample and for all subgroups. This means that the quantity $(1+\left[\frac{C}{100}\right]^2)$ was generally below 1.1, and the variation in sampling weights had little impact on the precision of sample estimates.

A few relatively large student weights were observed in some jurisdictions. An evaluation was made of the impact of trimming these largest weights back to a level consistent with the largest remaining weights found in the state. Such a procedure produced an appreciable reduction in the size of the coefficient of variation for these weights, and hence this trimming was implemented. Westat judged that this procedure had minimal potential to introduce bias, while the reduction in the coefficient of variation of the weights gives rise to an appreciable decrease in sampling error for the jurisdictions.

7.6 CALCULATION OF REPLICATE WEIGHTS

A replication method known as *jackknife* was used to estimate the variance of statistics derived from the full sample. The process of replication involves repeatedly selecting portions of the sample (replicates) and calculating the desired statistic (replicate estimates). The variability among the calculated replicate estimates is then used to obtain the variance of the full-sample estimate.

In each jurisdiction, replicates were formed in two steps. First, each school was assigned to one of a maximum of 62 replicate groups, each group containing at least one school. In the next step, a random subset of schools (or, in some cases, students within schools) in each replicate group was excluded. The remaining subset and all schools in the other replicate groups then constituted one of the 62 replicates. The process of forming these replicate groups, core to the process of variance estimation, is described below.

7.6.1 Defining Replicate Groups and Forming Replicates for Variance Estimation

Replicate groups were formed separately for public and nonpublic schools. Once replicate groups were formed for all schools, students were then assigned to their respective school replicate groups.

Public Schools. Noncertainty schools were sorted by jurisdiction according to sample type. Then within sample type, the schools were sorted by new school status and the order in which they were selected from the sampling frame. The schools were then grouped in pairs. Where there was an odd number of schools, the last replicate group contained three schools instead of two.

Each of the certainty public schools was assigned to one or more replicate groups of its own. If a school was selected three or more times in the sampling process, then it was assigned to two replicate groups. Here, schools were sorted by the estimated grade enrollment prior to group assignments. Depending on the jurisdiction, a maximum of 62 certainty groups were formed. The group numbering resumed from the last group number used for the noncertainty schools if the total number of public-school groups was less than 62. Otherwise, the numbering started from 62 down to the number needed for the last certainty public school. In the District of Columbia grade 8, which had only 36 certainty schools, the groups went from 1 to 53. Eighteen of the 36 certainty schools in the District of Columbia were selected three or more times and thus were assigned to two replicate groups. A replicate was formed by randomly deleting one half of the students in a certainty school from the sample. For certainty schools that were assigned to two replicate groups, the students were split equally between four "halves," two halves in each of the two replicate groups. This was repeated for each certainty school.

The purpose of this scheme was to assign as many replicates to a jurisdiction's public schools as permitted by the design, to a maximum of 62. When more than 62 replicates were assigned, the procedure ensured that no subset of the replicate groups (pairs of noncertainty schools, individual certainty schools, or groups of these) was substantially larger than the other replicate groups. The aim was to maximize the degrees of freedom available for estimating variances for public-school data.

A single replicate estimate was formed by dropping one member assigned to a particular replicate group. This process was repeated successively across replicate groups, giving up to 62 replicate estimates.

Nonpublic Schools. Replicate groups for noncertainty nonpublic schools were formed in one of the two methods described below. If any of the following conditions was true for a given jurisdiction, then the subsequent steps were taken to form replicate groups. Here, the numbering started at 62 down to the last needed number.

Conditions for Method 1:

- fewer than 11 nonpublic noncertainty schools; or
- fewer than 2 Catholic noncertainty schools; or
- fewer than 2 non-Catholic noncertainty schools.

Steps for Method 1:

- all schools were grouped into a single replicate group;
- schools were randomly sorted; and
- starting with the second school, replicates were formed by consecutively leaving out one of the remaining *n - 1* schools; each replicate included the first school.

When a given jurisdiction did not match conditions of the first method (i.e., when all of the following conditions were true), then the preceding steps were repeated separately for two groups, one consisting of Catholic schools and one consisting of non-Catholic schools.

Conditions for Method 2:

- more than 10 nonpublic noncertainty schools; and
- more than 1 Catholic noncertainty school; and
- more than 1 non-Catholic noncertainty school.

For jurisdictions with certainty nonpublic schools (Delaware, District of Columbia, Maryland, North Dakota, Rhode Island, Vermont, and Wyoming) each school was assigned to one or more groups. If a school was selected three or more times in the sampling, it was assigned to two groups. Prior to this assignment, schools were sorted in descending order of the estimated grade enrollment. The group numbering started at the last number where the noncertainty nonpublic schools ended. A replicate was formed by randomly deleting one half of the students in a certain school from the sample. For the certainty schools that were assigned to two replicate groups, the students were split equally between four “halves,” two halves in each of two replicate groups. This was repeated for each certainty school.

Again, the aim was to maximize the number of degrees of freedom for estimating sampling errors for nonpublic schools (and indeed for public and nonpublic schools combined) within the constraint of forming 62 replicate groups. Where a jurisdiction had a significant contribution from both Catholic and non-Catholic schools, Westat ensured that the sampling error estimates reflected the stratification on this characteristic.

Guam. For Guam, where all schools were selected with certainty, schools were assigned to one or more replicate groups proportional to their estimated grade enrollment.

DDESS and DoDDS Schools. Schools in the DDESS grade 8 sample were assigned to one or more replicate groups proportional to their estimated grade enrollment. Schools in the grade 4 DDESS and DoDDS samples were assigned to replicate groups following the general rules described above for all public schools.

7.6.2 School-Level Replicate Weights

As mentioned above, each replicate sample had to be reweighted to compensate for the dropped unit(s) defining the replicate. This reweighting was done in two stages. At the first-stage, the *i*th school included in a particular replicate *r* was assigned a replicate-specific school base weight defined as

$$W_{ri}^{sch} = K_r \times W_i^{sch}$$

where W_i^{sch} is the full-sample base weight for school *i*, and, for public schools

$$K_r = \begin{cases} 1.5 & \text{if school } i \text{ was contained in a "pair" consisting of 3 units from which} \\ & \text{the complementary member was dropped to form replicate } r, \\ 2 & \text{if school } i \text{ was contained in a pair consisting of 2 units from which the} \\ & \text{complementary member was dropped to form replicate } r, \\ 0 & \text{if school } i \text{ was dropped to form replicate } r, \text{ and} \\ 1 & \text{if school } i \text{ was not assigned to replicate } r \text{ or if school } i \text{ was a certainty.} \end{cases}$$

For nonpublic schools, Method 1:

$$K_r = \begin{cases} \dfrac{n}{n-1} & \text{if school } i \text{ was not dropped in forming replicate } r \\ 0 & \text{if school } i \text{ was dropped to form replicate } r \end{cases}$$

For nonpublic schools, Method 2 (with n_1 Catholic schools and n_2 non-Catholic schools):

$$K_r = \begin{cases} \dfrac{n_1}{n_1 - 1} & \text{if school } i \text{ was Catholic, not dropped from replicate } r, \\ & \text{and replicate } r \text{ was formed by dropping a Catholic school} \\ 1 & \text{if school } i \text{ was Catholic and replicate } r \text{ was formed by dropping a nonCatholic school} \\ \dfrac{n_2}{n_2 - 1} & \text{if school } i \text{ was nonCatholic, not dropped from replicate } r, \\ & \text{and replicate } r \text{ was formed by dropping a nonCatholic school} \\ 1 & \text{if school } i \text{ was nonCatholic and replicate } r \text{ was formed by dropping a Catholic school} \\ 0 & \text{if school } i \text{ was dropped to form replicate } r \end{cases}$$

Using the replicate-specific school base weights, W_{ri}^{sch}, the school-level nonresponse weighting adjustments were recalculated for each replicate *r*. That is, the school-level nonresponse adjustment factor for schools in replicate *r* and adjustment class *k* was computed as

$$F_{rk} = \frac{\sum_{i \in C_k} \left(W_{rki}^{sch} \times E_{ki} \right)}{\sum_{i \in C_k} \left(W_{rki}^{sch} \times E_{ki} \times \delta_{rki} \right)}$$

where

C_k = the subset of school records in adjustment class *k*;

W_{rki}^{sch} = the replicate-*r* base weight of the *i*th school in class *k*;

E_{ki} = the grade enrollment for the *i*th school in class *k*;

In the above formulation, the indicator variable δ_{rki} had a nonzero value only when the *i*th school in replicate *r* and adjustment class *k* participated in the assessment. The replicate-specific nonresponse-adjusted school weight for the *i*th school in replicate *r* in class *k* was then computed as

$$W_{rki}^{adj} = F_{rk} \times W_{rki}^{sch} \times \delta_{rki} .$$

7.6.3 Student-Level Replicate Weights

The replicate-specific adjusted student base weights were calculated by multiplying the replicate-specific adjusted school weights as described above by the corresponding within-school student weights. That is, the adjusted student base weight for the *j*th student in adjustment class *k* in replicate *r* was initially computed as

$$W_{rkij} = W_{rki}^{adj} \times W_{ij}^{within}$$

where

W_{rki}^{adj} = the nonresponse-adjusted school weight for school *i* in school adjustment class *k* and replicate *r*; and

W_{ij}^{within} = the within-school weight for the *j*th student in school *i*.

The final replicate-specific student weights were then obtained by applying the student nonresponse adjustment procedures to each set of replicate student weights. Let F_{rk} denote the student-level nonresponse adjustment factor for replicate *r* and adjustment class *k*. The final replicate-*r* student weight for student *j* in school *i* in adjustment class *k* was calculated as:

$$W_{rkij}^{final} = F_{rk} \times W_{rki}^{adj} \times W_{ij}^{within} .$$

Finally, estimates of the variance of sample-based estimates were calculated as

$$Var_{JK}(\hat{x}) = \sum_{r=1}^{62} (\hat{x}_r - \hat{x})^2 ,$$

where

$$\hat{x} = \sum_{i,j} W_{kij}^{final} \times x_{kij}$$

denotes an estimated total based on the full sample, and $\hat{x}_r$ denote the corresponding estimate based on replicate *r* with 62 replicates. The standard error of an estimate $\hat{x}$ is estimated by taking the square root of the estimated variance, $\mathrm{Var}_{JK}(\hat{x})$.

7.7 RAKING OF WEIGHTS

Raking (also known as iterative proportional fitting) is done in place of poststratification. Unlike poststratification, it is performed iteratively to two or more different distributions of a population total (i.e., gender and age). It is typically used in situations in which the interior cells of a cross-tabulation are either unknown, or some sample sizes in the cells are too small for efficient estimation. In raking, the marginal population totals, $N_{i.}$ and $N_{.j}$ are known (i.e., age and gender population counts), however, the interior cells of the cross-tabulation N_{ij} (the age by gender cells) are estimated from the sample by $\hat{N}_{ij}$, where these are the sum of weights in the cells. The raking algorithm proceeds by proportionally scaling the $\hat{N}_{ij}$, such that the following relations are satisfied:

$$\sum_{j} \hat{N}_{ij} = N_i$$

and

$$\sum_{i} \hat{N}_{ij} = N_j.$$

The 1996 State Assessment program used two different sets of inclusion rules indicated by sample type = 1 and sample type = 2. The science assessment was analyzed omitting the sample type = 1 SD/LEP students; and the mathematics assessment was analyzed omitting the sample type = 2 SD/LEP students. The SD/LEP student weights were raked separately for the two subsets as defined by sample type and public/nonpublic schools. Agreement was forced with totals estimated using both of the subsets combined for each of the school types. The purpose of this was to enhance the reliability (i.e., reduce the sampling error) of estimates produced by using information about student characteristics from the whole sample to enhance the estimates.

7.7.1 Raking Dimensions for Full Sample Student Weights

Public Schools. Five variables were used for the raking dimensions. These variables included two levels of SD (SD/nonSD), two levels of LEP (LEP/nonLEP), two levels of GENDER, five levels of RACE (White and Other; Black; Hispanic; Asian or Pacific Islander; and American Indian or Alaskan Native), and two levels of AGE. The variable AGE was defined as follows: for grade 4, those born in August 1985 or earlier and those born in September 1985 or later and for grade 8, those born in August 1981 or earlier and those born in September 1981 or later. Collapsing of levels was done so that no level of a single dimension contained fewer than 30 students for a state and grade.

Tables 7-10 and 7-11 show for each jurisdiction and for DDESS and DoDDS jurisdictions, the final collapsed levels that were used for the raking dimensions. A dash indicates that all levels were combined, and thus, the variable was not used as a raking dimension. An asterisk for the RACE variable indicates that all other levels of the dimension were combined into one level. For example in California, there are three levels of RACE: White and Other, Hispanic, and all others combined.

Nonpublic Schools. Because of the small numbers of nonpublic-school students, no raking was carried out. A factor of 2 was applied to the weights for the SD/LEP students since only half the SD/LEP sample was used for analysis.

7.7.2 Raking Student Replicate Weights

The replicate weights for the public SD/LEP students were raked similarly. Control totals for each replicate were calculated based on the totals for the replicate weights. The levels of the raking dimensions that were used for the replicates were the same collapsed levels as shown in Tables 7-10 and 7-11. For the nonpublic schools, again a factor of 2 was applied to the replicate weights of the SD/LEP students.

Table 7-10
Final Collapsed Levels Used for Raking Dimensions
DoDEA Jurisdictions, Grade 4

Other jurisdictions					
DoDEA/DDESS	-	-	-	-	-
DoDEA/DoDDS	Gender	Age	Race: W / H / *	-	-

Table 7-11

Final Collapsed Levels Used for Raking Dimensions
All Jurisdictions, Grade 8

Jurisdiction					
Alabama	Gender	Age	Race: W / *	-	-
Alaska	-	Age	Race: W / *	-	LEP
Arizona	Gender	Age	Race: H / *	SD	LEP
Arkansas	Gender	Age	Race: W / *	-	-
California	Gender	Age	Race: W / H / *	SD	LEP
Colorado	Gender	Age	Race: W / *	SD	-
Connecticut	Gender	Age	Race: W / *	-	LEP
Delaware	Gender	Age	Race: W / *	-	-
District of Columbia	Gender	Age	Race: B / *	-	-
Florida	Gender	Age	Race: W / H / *	SD	-
Georgia	Gender	Age	Race: W / *	-	-
Hawaii	Gender	-	Race: H / A / *	-	-
Indiana	Gender	Age	Race: W / *	-	-
Iowa	Gender	Age	-	-	-
Kentucky	Gender	Age	-	-	-
Louisiana	Gender	Age	Race: W / *	-	-
Maine	Gender	Age	-	-	-
Maryland	Gender	Age	Race: W / *	-	-
Massachusetts	Gender	Age	Race: W / H / *	-	-
Michigan	-	Age	-	-	-
Minnesota	Gender	Age	Race: W / *	-	-
Mississippi	Gender	-	Race: B / *	-	-
Missouri	Gender	Age	Race: W / *	-	-
Montana	Gender	Age	-	-	-
Nebraska	Gender	Age	Race: W / *	-	-
Nevada	-	-	-	-	-
New Hampshire	Gender	Age	-	-	-
New Jersey	Gender	Age	Race: W / *	-	-
New Mexico	Gender	Age	Race: W / H / *	SD	LEP
New York	Gender	Age	Race: W / H / *	SD	-
North Carolina	Gender	Age	Race: W / *	-	-
North Dakota	-	Age	-	-	-
Oregon	Gender	Age	Race: W / *	-	LEP
Rhode Island	Gender	Age	Race: W / *	SD	LEP
South Carolina	Gender	Age	Race: B / *	-	-
Tennessee	Gender	Age	Race: W / *	-	-
Texas	Gender	Age	Race: H / *	SD	LEP
Utah	Gender	-	Race: W / *	-	-
Vermont	Gender	Age	-	-	-
Virginia	Gender	Age	Race: W / *	-	-
Washington	Gender	Age	Race: W / *	SD	LEP
West Virginia	Gender	Age	-	-	-
Wisconsin	Gender	Age	Race: W / *	-	-
Wyoming	Gender	Age	Race: W / *	-	-
Other jurisdictions					
Guam	-	-	-	-	-
DoDEA/DDESS	-	-	-	-	-
DoDEA/DoDDS	Gender	-	Race: W / *	-	-

Chapter 8

THEORETICAL BACKGROUND AND PHILOSOPHY OF NAEP SCALING PROCEDURES[1]

Eugene G. Johnson and Nancy L. Allen
Educational Testing Service

8.1 OVERVIEW

The primary method by which results from the State Assessment are disseminated is scale-score reporting. With scaling methods, the performance of a sample of students in a subject area or subarea can be summarized on a single scale or a series of scales even when different students have been administered different items. This chapter presents an overview of the scaling methodologies employed in the analyses of the data from NAEP surveys in general and from the State Assessment in science in particular. Details of the scaling procedures specific to the 1996 State Assessment in science are presented in Chapter 9.

8.2 BACKGROUND

The basic information from an assessment consists of the responses of students to the items presented to them. For NAEP, these items are constructed to measure performance on sets of objectives developed by nationally representative panels of learning area specialists, educators, and concerned citizens. Satisfying the framework and specifications for the assessment and ensuring that the items selected to measure each part of the framework cover a range of difficulty levels typically requires many items. For example, the State Assessment in science required 195 items at grade 8 to meet the specifications provided for the assessment. To reduce student burden, each assessed student was presented only a fraction of the full pool of items through multiple matrix sampling procedures.

The most direct manner of presenting the assessment results is to report separate results for each item. However, because of the vast amount of information, having separate results for each of the items in the assessment pool hinders our understanding of the overall performance of subgroups of the population. Item-by-item reporting masks our understanding of similarities in trends and subgroup comparisons common across items.

An obvious way to summarize performance across a collection of items is to calculate the average of the separate item scores. The advantage of averaging is that it tends to cancel out the effects of peculiarities in items that can affect item difficulty in unpredictable ways. Furthermore, averaging makes it easier to compare the general performances of subpopulations.

Despite their advantages, there are a number of significant problems with average item scores. First, the interpretation of these averages depends on the items that happen to be

[1] Nancy L. Allen was responsible for the psychometric and statistical analysis of state and national NAEP data. Eugene G. Johnson is a senior psychometrician with special expertise in the design of NAEP and sampling issues. Robert J. Mislevy and Neal Thomas contributed greatly to previous versions of this chapter.

administered to a group of students. Since all students are not administered the same items, the average item score of students who happen to be administered a set of easy or difficult items would make that group's performance appear to be overly high or low. Second, again since the average score is related to the particular items administered direct comparisons of subpopulations become difficult because they require that those subpopulations be administered the same set of items. Third, because this approach limits comparisons to average scores on specific sets of items, it provides no simple way to report trends over time when the specific content of the item pool changes. Finally, direct estimates of quantities such as the proportion of students who would achieve a specific score across the items in the pool are not possible when every student is administered only a small fraction of the item pool. While the mean score across all items in the pool can be readily obtained (by calculating the average of the individual item scores), statistics that provide distributional scores across the full set of items in the pool cannot be readily obtained without additional assumptions.

These limitations can be overcome by the use of response scaling methods. When several items require similar skills, the regularities observed in response patterns can often be used to characterize both students and items using a relatively small number of variables. These variables include a student-specific variable, commonly called proficiency, estimated by the scale score, which quantifies a student's tendency to answer items correctly (or, for multipoint items, to achieve a certain score) and item-specific characteristics of an item such as its difficulty, effectiveness in distinguishing between students with different levels of proficiency, and chances of a very low proficiency student correctly answering a multiple-choice item. (These variables are discussed in more detail in the next section.) When combined through appropriate mathematical formulas, these characteristics capture the dominant features of the data. Furthermore, all students' proficiencies can be measured on a common scale, even though none of the students took all of the items in the pool. Using the common scale, it becomes possible to estimate distributions of proficiency in a population or subpopulation and to estimate the relationships between the scale scores and student background variables.

It is important to point out that any procedure of aggregation, whether it be a simple average or a more complex multidimensional scaling model, highlights certain patterns at the expense of other potentially interesting patterns that may reside within the data. Every item in a NAEP survey is of interest and can provide useful information about what young Americans know and can do. The choice of an aggregation procedure must be driven by a conception of which patterns are most important for a particular purpose.

The scaling for the State Assessment in science was carried out separately for each of the three fields of science specified in the framework for grade 8 science. This scaling within subareas was done because it was anticipated that different patterns of performance might exist for these essential subdivisions of the subject area. The three fields of science are: *earth science*, *physical science*, and *life science*. By creating a separate scale for each field of science, any differences in subpopulation performance between the fields are preserved.

The creation of a series of separate scales to describe science performance does not preclude the reporting of a single index of overall science performance — that is, an overall science composite. A composite is computed as the weighted average of the three fields of science scales, where the weights correspond to the relative importance given to each field as defined by the framework. The composite provides a global measure of performance within science, whereas the fields of science scales allow the measurement of important interactions among these subdivisions of science.

8.3 SCALING METHODOLOGY

This section reviews the scaling models employed in the analyses of data from the State Assessment in science and the 1996 national science assessment. It also reviews the multiple imputation or "plausible values" methodology that allows such models to be used with NAEP's sparse item-sampling design. The reader is referred to Mislevy (1991) for an introduction to plausible values methods and a comparison with standard psychometric analyses, to Mislevy, Johnson, and Muraki (1992) and Beaton and Johnson (1992) for additional information on how the models are used in NAEP, and to Rubin (1987) for the theoretical underpinnings of the approach. It should be noted that the imputation procedure used by NAEP is a mechanism for providing plausible values for scale score averages and not for filling in blank responses to background or cognitive variables.

While the NAEP procedures were developed explicitly to handle the characteristics of NAEP data, they build on work paralleled by other researchers. See, for example Dempster, Laird, and Rubin (1977); Little and Rubin (1983, 1987); Andersen (1980); Engelen (1987); Hoijtink (1991); Laird (1978); Lindsey, Clogg, and Grego (1991); Zwinderman (1991); Tanner and Wong (1987); and Rubin (1987, 1991).

The 195 items administered at grade 8 in the State Assessment were also administered to students of the same grades in the national science assessment. The number of items actually scaled differs from the number of items administered due to decisions about the treatment of items in scaling (see Tables 9-1 through 9-4 in Chapter 9). However, because the administration procedures differed, the State Assessment data were scaled independently from the national data. The national assessment also included results for students in grades 4 and 12. Details of the scaling of the State Assessment and the subsequent linking to the results from the national science assessment are provided in Chapter 9.

8.3.1 The Scaling Models

Three distinct scaling models, depending on item type and scoring procedure, were used in the analysis of the data from the State Assessment. Each of the models is based on item response theory (IRT; e.g., Lord, 1980). Each is a "latent variable" model, which is defined separately for each of the scales at each grade. A latent variable model expresses students' tendencies to respond (such as correct/incorrect) on the items as a function of a characteristic that is not directly observed. This characteristic is called proficiency. Students' proficiencies are estimated by scale scores.

A three-parameter logistic (3PL) model was used for the multiple-choice items (which were scored correct/incorrect). The fundamental equation of the 3PL model is the probability that a student, whose proficiency on scale *k* is characterized by the *unobservable* variable θ_k, will respond correctly to item *j*:

$$P(x_j = 1 | \theta_k, a_j, b_j, c_j) = c_j \frac{(1 - c_j)}{1 + \exp[-1.7 a_j(\theta_k - b_j)]} \equiv P_{j1}(\theta_k) \tag{8.1}$$

In Equation 8.1:

x_j is the response to item *j*, 1 if correct and 0 if not;

a_j where $a_j > 0$, is the slope parameter of item *j*, characterizing the strength of its relationship to the latent proficiency;

b_j is the location parameter of item *j*, characterizing its difficulty with respect to the latent proficiency; and

c_j where $0 \le c_j < 1$, is the lower asymptote parameter of item *j*, reflecting the chances of students of very low proficiency selecting the correct option.

Further, the probability of an incorrect response to the item is defined as:

$$P_{j0} = P(x_j = 0 | \theta_k, a_j, b_j, c_j) = 1 - P_{j1}(\theta_k). \tag{8.2}$$

A two-parameter logistic (2PL) model was used for short constructed-response items, which were scored correct or incorrect. The equations of the 2PL model are the same as those of Equations 8.1 and 8.2, with the c_j parameter fixed at zero.

In addition to the 75 multiple-choice and 6 short constructed-response items, 105 extended constructed-response items were administered in the grade 8 State and national assessments. Each of these items was scored on a multipoint ranging from 0 to 3 or 0 to 4. As a result of examining the responses to these items, some of the response categories for several of these items were combined for inclusion in the final IRT scales. Additionally, as discussed in Chapter 9, certain sets of items consisting of highly correlated parts were combined into four cluster items or "testlets" (Wainer & Kiely, 1987) where the score assigned to a cluster item was the number of constituent parts answered correctly. (See Chapter 9 for a description of the special treatment of items and item categories during scaling.) Items that are scored on a multipoint scale are referred to as polytomous items. The multiple-choice and short constructed-response items, which are scored correct/incorrect, are referred to as dichotomous items.

The polytomous items were scaled using a generalized partial credit model (Muraki, 1992). The fundamental equation of this model is the probability that a person, whose proficiency on scale *k* is characterized by the unobservable variable θ_k, will respond to item *j* in a way to be scored *i*:

$$P(x_j = i | \theta_k, a_j, b_j, d_{j,1}, \ldots, d_{j,m_j-1}) = \frac{\exp\left(\sum_{v=0}^{i} 1.7 a_j(\theta_k - b_j + d_{j,v})\right)}{\sum_{g=0}^{m_j-1} \exp\left(\sum_{v=0}^{g} 1.7 a_j(\theta_k - b_j + d_{j,v})\right)} \equiv P_{ji}(\theta_k) \tag{8.3}$$

where

m_j is the number of ordered categories in response to item *j*

x_j is the response to item j, with possibilities $i = 0,1,...,m_j - 1$

a_j is the slope parameter;

b_j is the item location parameter characterizing overall difficulty with respect to the latent proficiency; and

$d_{j,i}$ is the category i threshold parameter (see below).

Indeterminacies in the parameters of the above model are resolved by setting $d_{j,0} = 0$ and setting $\sum_{i=1}^{m_j-1} d_{j,i} = 0$. Muraki (1992) points out that $b_j - d_{j,i}$ is the point on the θ_k scale at which the plots of $P_{j,i-1}(\theta_k)$ and $P_{ji}(\theta_k)$ intersect and so characterizes the point on the θ_k scale at which the student's response to item j has equal probability of falling in score category $i - 1$ and falling in score category i.

When $m_j = 2$, so that there are two score categories (such as 0,1), it can be shown that $P_{ji}(\theta_k)$ of Equation 8.3 for i=0,1 corresponds respectively to $P_{j0}(\theta_k)$ and $P_{j1}(\theta_k)$ of the 2PL model (Equations 8.1 and 8.2 with c_j=0).

A typical assumption of item response theory is the independence of the response by a student to a set of items, given or conditional on the student's proficiency. That is, for a student with a specific proficiency of θ_k, the joint probability of a particular response pattern $\underline{x} = (x_1,...,x_n)$ across a set of n items is simply the product of terms based on Equations 8.1, 8.2, and 8.3:

$$P(\underline{x} \mid \theta_k, item\ parameters) = \prod_{j=1}^{n} \prod_{i=0}^{m_j-1} P_{ji}(\theta_k)^{u_{ji}} \tag{8.4}$$

where $P_{ji}(\theta_k)$ is of the form appropriate to the type of item (dichotomous or polytomous), m_j is taken equal to 2 for the dichotomously scored items, and u_{ji} is an indicator variable defined by

$$u_{ji} = \begin{cases} 1 \text{ if response } x_j \text{ was in score category } i \\ 0 \text{ otherwise} \end{cases}$$

It is also typically assumed that response probabilities are conditionally independent of background variables $(\underline{y})$, given θ_k, or

$$P(\underline{x} \mid \theta_k, item\ parameters, \underline{y}) = p(\underline{x} \mid \theta_k, item\ parameters) \tag{8.5}$$

After $\underline{x}$ has been observed, Equation 8.4 can be viewed as a likelihood function, and provides a basis for inference about θ_k or about item parameters. Estimates of item parameters (see Appendix D) were obtained by the NAEP BILOG/PARSCALE program, which combines Mislevy and Bock's (1982) BILOG and Muraki and Bock's (1991) PARSCALE computer programs, and which concurrently estimates parameters for all items (dichotomous and polytomous). The item parameters are then treated as known in subsequent calculations. The parameters of the items constituting each of the separate scales were estimated independently of the parameters of the other scales. Once items have been calibrated in this manner, a likelihood function for the scale proficiency θ_k is induced by a vector of responses to any subset of calibrated items, thus allowing θ_k-based inferences from matrix samples.

As stated previously, item parameter estimation was performed independently for the State Assessment and for the national science assessment. In both cases, the identical scale definitions were used.

In all NAEP IRT analyses, missing responses *at the end of each block* of items a student was administered were considered "not-reached," and treated as if they had not been presented to the student. Missing responses to dichotomous items *before the last observed response in a block* were considered intentional omissions, and treated as fractionally correct and assigned a score equal to the reciprocal of the number of response alternatives. These conventions are discussed by Mislevy and Wu (1988). With regard to the handling of not-reached items, Mislevy and Wu found that ignoring not-reached items introduces slight biases into item parameter estimation. The degree of this bias depends on the number of not-reached items and whether speed is correlated with ability. With regard to omissions, they found that the method described above provides consistent limited-information likelihood estimates of item and ability parameters under the assumption that students omit only if they can do no better than responding randomly.

Although the IRT models are employed in NAEP only to summarize performance, a number of checks are made to detect serious violations of the assumptions underlying the models (such as conditional independence). When warranted, remedial efforts are made to mitigate the effects of such violations on inferences. These checks include comparisons of empirical and theoretical item response functions to identify items for which the IRT model may provide a poor fit to the data.

The scales in NAEP are determined *a priori* by grouping items into the fields of science defined by the frameworks developed by the National Assessment Governing Board. A proficiency scale θ_k is defined *a priori* by the collection of items representing each scale. What is important, therefore, is that the models capture salient information in the response data to effectively summarize the overall performance on each field of the populations and subpopulations being assessed in the fields of science. NAEP routinely conducts differential item functioning (DIF) analyses to guard against potential biases in making subpopulation comparisons based on the scale score distributions.

The local independence assumption embodied in Equation 8.4 implies that item response probabilities depend only on θ_k and the specified item parameters, and not on the position of the item in the booklet, the content of other items near an item of interest, the test-administration conditions, or the timing conditions. However, these factors are certainly present in any administration. The practical question is whether inferences based on the IRT probabilities obtained via Equation 8.4 are robust with respect to these violations of the ideal assumptions underlying the IRT model. Our experience with the 1986 NAEP reading anomaly (Beaton & Zwick, 1990) has shown that for measuring small changes over time, changes in item context and speededness conditions can lead to unacceptably large error components. These can be avoided by presenting items used to measure change in identical test forms, with identical timings and administration conditions. Thus, we do *not* maintain that the item parameter estimates obtained in any particular booklet configuration are appropriate for other conceivable booklet configurations. Rather, we assume that the parameter estimates are context-bound. (For this reason, we prefer common population equating to common item equating whenever equivalent random samples are available for linking.) This is the reason that the data from the State Assessment were calibrated separately from the data from the national NAEP — since the administration procedures differed somewhat between the State Assessment and the national NAEP, the values of the item parameters could be different. Chapter 9 provides details on the procedures used to link the results of the 1996 State Assessment to those of the 1996 national assessment.

8.3.2 An Overview of Plausible Values Methodology

Item response theory was developed in the context of measuring individual students' performance. In that setting, each student is administered enough items (often 60 or more) to permit precise estimation of their latent proficiency θ. This may be accomplished by using a maximum likelihood estimate $\hat{\theta}$, for example. When there are enough items administered to each student, the uncertainty associated with each θ is negligible. As a result, the distribution of θ, or the joint distribution of θ with other variables, can then be closely approximated using students' $\hat{\theta}$ values as if they were true θ values.

This approach breaks down in the assessment setting when, in order to provide broader content coverage in limited testing time, each student is administered relatively few items in a scale. The problem is that the uncertainty associated with individual proficiencies is too large to ignore, and the $\hat{\theta}$ distribution, as an estimate of the θ distribution, can be seriously biased. (The failure of this approach was verified in early analyses of the 1984 NAEP reading survey; see Wingersky, Kaplan, & Beaton, 1987.) Plausible value methodology was developed as a way to estimate key population features consistently. A detailed development of plausible values methodology is given in Mislevy (1991). Along with theoretical justifications, that paper presents comparisons with standard procedures, discussions of biases that arise in some secondary analyses, and numerical examples.

The following provides a brief overview of the plausible values approach, focusing on its implementation in the State Assessment analyses.

Let $\underline{y}$ represent the responses of all sampled examinees to background and attitude questions, along with design variables such as school membership, and let $\underline{\theta}$ represent the vector of scale proficiency values. If θ were known for all sampled examinees, it would be possible to compute a statistic $t(\underline{\theta},\underline{y})$ — such as a scale or composite subpopulation sample mean, a sample percentile point, or a sample regression coefficient — to estimate a corresponding population quantity T. A function $U(\underline{\theta},\underline{y})$ — e.g., a jackknife estimate — would be used to gauge sampling uncertainty, as the variance of t around T in repeated samples from the population.

Because the scaling models are latent variable models, however, θ values are not observed even for sampled students. To overcome this problem, we follow Rubin (1987) by considering $\underline{\theta}$ as "missing data" and approximate $t(\underline{\theta},\underline{y})$ by its expectation given $(\underline{x},\underline{y})$, the data that actually were observed, as follows:

$$\begin{aligned} t^*(\underline{x},\underline{y}) &= E[t(\underline{\theta},\underline{y})|\underline{x},\underline{y}] \\ &= \int t(\underline{\theta},\underline{y})\, p(\underline{\theta}\,|\underline{x},\underline{y})\, d\underline{\theta}\,. \end{aligned} \tag{8.6}$$

It is possible to approximate t^* using random draws from the conditional distribution of the scale score averages given the item responses x_i, background variables y_i, and model parameters for sampled student i. These values are referred to as imputations in the sampling literature, and plausible values in NAEP. The value of θ for any student that would enter into the computation of t is thus replaced by a randomly selected value from the student's conditional distribution. Rubin (1987) proposes that this process be carried out several times — multiple imputations — so that the uncertainty associated with imputation can be quantified. The average of the results of, M estimates of t, each computed from a different set of plausible values, is a Monte Carlo approximation of Equation 8.6; the variance among them, B, reflects uncertainty due to not observing $\underline{\theta}$, and must be added to the estimated expectation of $U(\underline{\theta},\underline{y})$, which reflects uncertainty due to testing only a sample of students from the population. Section 8.5 explains how plausible values are used in subsequent analyses.

It cannot be emphasized too strongly that **plausible values are *not* test scores for *individuals*** in the usual sense. Plausible values are offered only as intermediary computations for calculating integrals of the form of Equation 8.6, in order to estimate *population* characteristics. When the underlying model is correctly specified, plausible values will provide consistent estimates of population characteristics, even though they are not generally unbiased estimates of the scale score averages of the individuals with whom they are associated. The key idea lies in a contrast between plausible values and the more familiar θ estimates of educational measurement that are in some sense optimal for each examinee (e.g., maximum likelihood estimates, which are consistent estimates of an examinee's θ, and Bayes estimates, which provide minimum mean-squared errors with respect to a reference population): *Point estimates that are optimal for individual examinees have distributions that can produce decidedly nonoptimal (specifically, inconsistent) estimates of population characteristics* (Little & Rubin, 1983). Plausible values, on the other hand, are constructed explicitly to provide consistent estimates of population effects. For further discussion see Mislevy, Beaton, Kaplan, and Sheehan (1992).

8.3.3 Computing Plausible Values in IRT-based Scales

Plausible values for each student r are drawn from the conditional distribution $p(\underline{\theta}_r|\underline{x}_r,\underline{y}_r,\Gamma,\Sigma)$, where Γ and Σ are regression model parameters defined in this subsection. This subsection describes how, in IRT-based scales, these conditional distributions are characterized, and how the draws are taken. An application of Bayes' theorem with the IRT assumption of conditional independence produces

$$p(\underline{\theta}_r|\underline{x}_r,\underline{y}_r,\Gamma,\Sigma) \propto P(\underline{x}_r|\underline{\theta}_r,\underline{y}_r,\Gamma,\Sigma)\, p(\underline{\theta}_r|\underline{y}_r,\Gamma,\Sigma) = P(\underline{x}_r|\underline{\theta}_r)\, p(\underline{\theta}_r|\underline{y}_r,\Gamma,\Sigma)\ , \tag{8.7}$$

where, for vector-valued $\underline{\theta}_r$, $P(\underline{x}_r|\underline{\theta}_r)$ is the product over scales of the *independent likelihoods* induced by responses to items within each scale, and $p(\underline{\theta}_r|\underline{y}_r,\Gamma,\Sigma)$ is the multivariate — and generally nonindependent — *joint density* of scale score averages for the scales, conditional on the observed value $\underline{y}_r$ of a student's background responses, and the parameters Γ and Σ. The scales are determined by the item parameter estimates that constrain the population mean to zero and standard deviation to one. The item parameter estimates are fixed and regarded as population values in the computation described in this subsection.

In the analyses of the data from the State Assessment and the data from the national science assessment, a normal (Gaussian) form was assumed for $p(\underline{\theta}_r|\underline{y}_r, \Gamma, \Sigma)$, with a variance-covariance matrix, Σ, and with a mean given by a linear model with slope parameters, Γ, based on the first 134 to 200 principal components of 482 selected main effects and two-way interactions of the complete vector of background variables. The variance-covariance matrix, Σ, is common across different patterns of responses to the background variable $\underline{y}$. The included principal components will be referred to as the *conditioning variables*, and will be denoted y^c. (The complete set of original background variables used in the State Assessment science analyses are listed in Appendix C.) The following model was fit to the data within each jurisdiction:

$$\underline{\theta} = \Gamma' \underline{y}^c + \underline{\varepsilon}, \tag{8.8}$$

where $\underline{\varepsilon}$ is multivariately normally distributed with mean zero and variance-covariance matrix Σ. The number of principal components of the conditioning variables used for each jurisdiction was sufficient to account for 90 percent of the total variance of the full set of conditioning variables (after standardizing each variable). As in regression analysis, Γ is a matrix each of whose columns contains the *effects* for one scale and Σ is the *variance-covariance matrix of residuals* between scales. By fitting the model (Equation 8.8) separately within each jurisdiction, interactions between each jurisdiction and the conditioning variables are automatically included in the conditional joint density of scale score averages.

Maximum likelihood estimates of Γ and Σ, denoted by $\hat{\Gamma}$ and $\hat{\Sigma}$, are obtained from an enhancement of Sheehan's (1985) MGROUP computer program using the EM algorithm described in Mislevy (1985). The enhanced version is referred to as CGROUP (Thomas, 1993). The EM algorithm requires the computation of the mean, $\underline{\bar{\theta}}_r$, and variance, Σ_r^p, of the posterior distribution in Equation 8.7. These moments are computed using higher order asymptotic corrections (Thomas, 1993).

After completion of the EM algorithm, the plausible values are drawn in a three-step process from the joint distribution of the values of Γ for all sampled students. First, a value of Γ is drawn from a normal approximation to $P(\Gamma, \Sigma \mid x_r, y_r)$ that fixes Σ at the value $\hat{\Sigma}$, (Thomas, 1993). Second, conditional on the generated value of Γ (and the fixed value of $\Sigma = \hat{\Sigma}$), the mean, $\underline{\bar{\theta}}_r$, and variance, Σ_r^p, of the posterior distribution in Equation 8.7 (i.e., $p(\underline{\theta}_r|x_r, y_r, \Gamma, \Sigma)$) are computed using the same methods applied in the EM algorithm. In the third step, the $\underline{\theta}_r$ are drawn independently from a multivariate normal distribution with mean $\underline{\bar{\theta}}_r$ and variance Σ_r^p, approximating the distribution in Equation 8.7. These three steps are repeated five times producing five imputations or plausible values of $\underline{\bar{\theta}}_r$ for each sampled student.

8.4 ANALYSES

When survey variables are observed without error from every student, standard variance estimators quantify the uncertainty associated with sample statistics from the only source of uncertainty, namely the sampling of students. Item-level statistics for NAEP cognitive items meet this requirement, but scale-score values, which estimate proficiency, do not. The IRT models used in their construction posit an unobservable proficiency variable θ to summarize performance on the items in the subarea. The fact that θ values are not observed even for the students in the sample requires additional statistical analyses to draw inferences about θ distributions and to quantify the uncertainty associated with those inferences. As described above, Rubin's (1987) multiple imputations procedures were adapted to the context of latent variable models to produce the plausible values upon which many analyses of the data from the State Assessment were based. This section describes how plausible values were employed in subsequent analyses to yield inferences about population and subpopulation distributions of scale score averages.

8.4.1 Computational Procedures

Even though one does not observe the θ value of student r, one does observe variables that are related to it: $\underline{x}_r$, the student's answers to the cognitive items he or she was administered in the area of interest, and $\underline{y}_r$, the student's answers to demographic and background variables. Suppose one wishes to draw inferences about a number $T(\underline{\theta},\underline{Y})$ that could be calculated explicitly if the θ and $\underline{y}$ values of each member of the population were known. Suppose further that if θ values were observable, we would be able to estimate T from a sample of N pairs of θ and $\underline{y}$ values by the statistic $t(\underline{\theta},\underline{y})$ [where $(\underline{\theta},\underline{y}) \equiv (\theta_1,\underline{y}_1,\ldots,\theta_N,\underline{y}_N)$], and that we could estimate the variance in t around T due to sampling students by the function $U(\underline{\theta},\underline{y})$. Given that observations consist of $(\underline{x}_r,\underline{y}_r)$ rather than $(\underline{\theta}_r,\underline{y}_r)$, we can approximate t by its expected value conditional on $(\underline{x},\underline{y})$, or

$$t^*(\underline{x},\underline{y}) = E[t(\underline{\theta},\underline{y})|\underline{x},\underline{y}] = \int t(\underline{\theta},\underline{y})\, p(\underline{\theta}\;|\underline{x},\underline{y})\, d\underline{\theta}\,.$$

It is possible to approximate t^* with random draws from the conditional distributions $p(\underline{\theta}_r|\underline{x}_r,\underline{y}_r)$, which are obtained for all students by the method described in Section 8.3.3. Let $\underline{\hat{\theta}}_m$ be the mth such vector of plausible values, consisting of a multidimensional value for the latent variable of each student. This vector is a plausible representation of what the true $\underline{\theta}$ vector might have been, had we been able to observe it.

The following steps describe how an estimate of a scalar statistic $t(\underline{\theta},\underline{y})$ and its sampling variance can be obtained from M (>1) such sets of plausible values. (Five sets of plausible values are used in analyses of the NAEP State Assessment.)

1. Using each set of plausible values $\underline{\hat{\theta}}_m$ in turn, evaluate t as if the plausible values were true values of $\underline{\theta}$. Denote the results $\hat{t}_m$ for m=1,...,M.

2. Using the jackknife variance estimator defined in Chapter 7, compute the estimated sampling variance of $\hat{t}_m$, denoting the result U_m.

3. The final estimate of t is

$$t^* = \sum_{m=1}^{M} \frac{\hat{t}_m}{M}.$$

4. Compute the average sampling variance over the M sets of plausible values, to approximate uncertainty due to sampling students:

$$U^* = \sum_{m=1}^{M} \frac{U_m}{M}.$$

5. Compute the variance among the M estimates $\hat{t}_m$, to approximate uncertainty due to not observing θ values from students:

$$B = \sum_{m=1}^{M} \frac{(\hat{t}_m - t^*)^2}{(M - 1)}$$

6. The final estimate of the variance of t^* is the sum of two components:

$$V = U^* + (1 + M^{-1})\, B.$$

Note: Due to the excessive computation that would be required, NAEP analyses did not compute and average jackknife variances over all five sets of plausible values, but only on the first set. Thus, in NAEP reports, U^* *is approximated by* U_1.

8.4.2 Statistical Tests

Suppose that if θ values were observed for sampled students, the statistic $(t - T)/U^{1/2}$ would follow a t-distribution with d degrees of freedom. Then the incomplete-data statistic $(t^* - T)/V^{1/2}$ is approximately t-distributed, with degrees of freedom given by

$$\nu = \frac{1}{\dfrac{f^2}{M - 1} + \dfrac{(1 - f)^2}{d}}$$

where f is the proportion of total variance due to not observing θ values:

$$f = (1+M^{-1})\, B / V.$$

When B is small relative to U^*, the reference distribution for incomplete-data statistics differs little from the reference distribution for the corresponding complete-data statistics. This is the case with main NAEP reporting variables. If, in addition, d is large, the normal approximation can be used to flag “significant” results.

For k-dimensional t, such as the k coefficients in a multiple regression analysis, each U_m and U^* is a covariance matrix, and B is an average of squares and cross-products rather than simply an average of squares. In this case, the quantity $(T\text{-}t^*)\, V^{-1}\, (T\text{-}t^*)'$ is approximately F distributed, with degrees of freedom equal to k and ν, with ν defined as above but with a matrix generalization of f:

$$f = (1+M^{-1})\, \mathrm{Trace}\, (BV^{-1})/k\, .$$

By the same reasoning as used for the normal approximation for scalar t, a chi-square distribution on k degrees of freedom often suffices.

8.4.3 Biases in Secondary Analyses

Statistics t^* that involve scale score averages in a scaled field of science and variables included in the conditioning variables y^c are consistent estimates of the corresponding population values *T*. Statistics involving background variables *y* that were *not* conditioned on, or relationships among scale score averages from *different* fields of science, are subject to asymptotic biases whose magnitudes depend on the type of statistic and the strength of the relationships of the nonconditioned background variables to the variables that were conditioned on and to the score scale of interest. That is, the large sample expectations of certain sample statistics need not equal the true population parameters.

The *direction* of the bias is typically to underestimate the effect of nonconditioned variables. For details and derivations see Beaton and Johnson (1990), Mislevy (1991), and Mislevy and Sheehan (1987, Section 10.3.5). For a given statistic t^* involving one field of science and one or more nonconditioned background variables, the *magnitude* of the bias is related to the extent to which observed responses, *x*, account for the latent variable θ, and the degree to which the nonconditioned background variables are correlated with the conditioning background variables. The first factor — conceptually related to test reliability — acts consistently in that greater measurement precision reduces biases in *all* secondary analyses. The second factor acts to reduce biases in certain analyses but increase it in others. In particular,

- High shared variance between conditioned and nonconditioned background variables *mitigates* biases in analyses that involve only scale score and nonconditioned variables, such as marginal means or regressions.

- High shared variance *exacerbates* biases in regression coefficients of conditional effects for nonconditioned variables, when nonconditioned and conditioned background variables are analyzed jointly as in multiple regression.

The large number of background variables that have been included in the conditioning vector for the State Assessment allows a large number of secondary analyses to be carried out with little or no bias, and mitigates biases in analyses of the marginal distributions of θ in nonconditioned variables. Kaplan and Nelson's analysis of the 1988 NAEP reading data (some results of which are summarized in Mislevy, 1991), which had a similar design and fewer conditioning variables, indicates that the potential bias for nonconditioned variables in multiple regression analyses is below 10 percent, and biases in simple regression of such variables is below 5 percent. Additional research (summarized in Mislevy, 1990) indicates that most of the bias reduction obtainable from conditioning on a large number of variables can be captured by instead conditioning on the first several principal components of the matrix of all original conditioning variables. This procedure was adopted for the State Assessment by replacing the conditioning effects by the first *K* principal components, where *K* was selected so that 90 percent of the total variance of the full set of conditioning variables (after standardization) was captured. Mislevy (1990) shows that this puts an upper bound of 10 percent on the average bias for all analyses involving the original conditioning variables.

Chapter 9

DATA ANALYSIS AND SCALING FOR THE 1996 STATE ASSESSMENT PROGRAM IN SCIENCE[1]

Spencer S. Swinton, John R. Donoghue, Steven P. Isham,
Lois H. Worthington, and Ingeborg U. Novatkoski
Educational Testing Service

9.1 OVERVIEW

This chapter describes the analyses used in developing the 1996 State Assessment science scales. The procedures used were similar to those employed in the analysis of the 1990, 1992, and 1996 State Assessments in mathematics, and the 1992 and 1994 State Assessments in reading, and are based on the philosophical and theoretical rationale given in the previous chapter. For 1996, the NAEP science assessment framework incorporated a balance of knowledge and skills based on current reform reports, exemplary curriculum guides, and research on the teaching and learning of science. The 1996 State Assessment included the assessment of both public- and nonpublic-school students for most jurisdictions and contained hands-on science tasks and theme blocks as well as considerably more constructed-response items.

There were five major steps in the analysis of the State Assessment science data, each of which is described in a separate section:

- conventional item and test analyses, and DIF analyses (Section 9.3);
- item response theory (IRT) scaling (Section 9.4);
- estimation of state and subgroup proficiency distributions based on the "plausible values" methodology (Section 9.5);
- linking of the 1996 State Assessment scales to the corresponding scales from the 1996 national assessment (Section 9.6); and
- creation of the State Assessment science composite scale (Section 9.7).

Section 9.8 provides an explanation of sampling weights, and Section 9.9 discusses the special grade 4 assessment of Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) and Department of Defense Dependents Schools (DoDDS).

To set the context within which to describe the methods and results of scaling procedures, a brief review of the assessment instruments and administration procedures is provided.

[1] Spence Swinton had the primary responsibility for the planning and coordination of the science state analyses. John Donoghue was responsible for the planning and coordination of the science national analyses and assisted with the science state analyses. Computer activities for these analyses were directed by Steve Isham and completed by Lois Worthington and Inge Novatkoski.

9.2 DESCRIPTION OF ITEMS, ASSESSMENT BOOKLETS, AND ADMINISTRATION PROCEDURES

The general design structure of the 1996 State Assessment in science was similar to those of previous State Assessments. However, the particulars of the 1996 design differed in several respects from earlier designs. Prior to the 1994 Trial State Assessment in reading (Mazzeo, Allen, & Kline, 1995), earlier assessments were administered to public-school students only. The 1994 and 1996 assessments included samples of both public- and nonpublic-school students. In the 1996 State Assessment, an attempt was made to include more students with disabilities (SD) and students with limited English proficiency (LEP) by liberalizing inclusion rules. Further, a special study of fourth-grade students enrolled in DDESS and DoDDS was completed. Although data from the Department of Defense Education Activity (DoDEA) grade 8 students were treated in the same manner as those from any other jurisdiction in the State Assessment, the grade 4 students from DoDEA schools were administered the national assessment grade 4 materials, and their mean scale scores were reported on the national scales. Information concerning the analysis of DoDEA grade 4 students can be found in Section 9.9 of this chapter, and in a separate DoDEA summary report.

The eighth-grade item pool was identical to that of the national science assessment, containing 190 items (after scaling), each of which was classified into one of three fields of science scales: 63 items for *earth science*; 62 items for *physical science*; and 65 items for *life science*. These items consisted of 75 multiple-choice items, 6 constructed-response items scored dichotomously, 105 constructed-response items scored polytomously, and 4 "cluster items."[2] The items were divided into 15 mutually exclusive blocks. Each student's booklet contained three blocks of cognitive items. Four of the blocks were hands-on tasks in which students were given a set of equipment and asked to conduct an investigation and answer questions (mostly constructed-response) relating to the investigation. These hands-on tasks were presented in the last position in every booklet, after two paper-and-pencil blocks. Three of the paper-and-pencil blocks were theme blocks. Theme blocks were placed randomly in the second position of some of the student booklets, but not in every booklet. No student received more than one theme block. Each theme block was paired with each non-theme paper-and-pencil block just once. Each paper-and-pencil block appeared in the first or second position the same number (3 or 4) of times. The composition of each block of items, in terms of content and format, is given in Table 9-1.[3] Some items (fewer than 10%) received special treatment during scaling. Table 9-2 shows the composition of each block after deletions of items and collapsing of categories for polytomously-scored constructed-response items as a result of scaling. If data had poor fit with the response model for an item, the item was deleted. If a constructed-response item was scored in multiple categories but one category had no (or very few) responses or had responses that did not fit the data in some categories, it was combined with other categories. All item deletions and common decisions about the treatment of items in scaling were made for the State and national assessments in science.

[2] A cluster item is an aggregation of a group of items (in the case of NAEP science, typically two to four items) that are related to a single content strand, topic, or stimulus, and are developed and scored as a single unit (see Wainer & Kiely, 1987, for further details and examples of different types of cluster items). Some items were initially scored as cluster items, and the additional clusters were formed during scaling.

[3] The two cluster items in Table 9-1 were scored as cluster items before scaling. Two additional cluster items (shown in Table 9-2) were formed during scaling.

Table 9-1

*1996 NAEP Science Block Composition by Fields of Science Scale and Item Type**

(As defined before scaling.)

Block	Multiple-Choice Items	Constructed-Response Items Scored Dichotomously	Constructed-Response Items Scored Polytomously	Cluster Items	Total Items
S3	0	0	5	1	6
S4	3	0	7	0	10
S5	0	0	8	0	8
S6	0	0	7	0	7
S7	2	2	8	0	12
S8	5	0	5	0	10
S9	3	0	10	0	13
S10	8	1	7	0	16
S11	8	0	8	0	16
S12	8	1	7	0	16
S13	8	0	8	0	16
S14	7	0	8	1	16
S15	7	1	8	0	16
S20	8	0	8	0	16
S21	7	0	9	0	16
Total	**74**	**5**	**113**	**2**	**194**

Table 9-2

*1996 NAEP Science Block Composition by Fields of Science Scale and Item Type**

(As defined after scaling. Counts reflect items that were dropped and collapsed.)

Block	Multiple-Choice Items	Constructed-Response Items Scored Dichotomously	Constructed-Response Items Scored Polytomously	Cluster Items	Total Items
S3	0	1	4	1	6
S4	2	1	5	1	9
S5	0	1	7	0	8
S6	0	1	5	0	6
S7	2	4	6	0	12
S8	5	1	4	0	10
S9	3	2	8	0	13
S10	8	1	7	0	16
S11	8	1	7	0	16
S12	8	1	7	0	16
S13	8	1	5	1	15
S14	7	0	8	1	16
S15	6	1	8	0	15
S20	8	2	6	0	16
S21	7	0	9	0	16
Total	**72**	**18**	**96**	**4**	**190**

The 11 paper-and-pencil blocks contained from 5 to 10 constructed-response items. The four hands-on task blocks contained from six to eight constructed-response items. One of the hands-on task blocks also contained three multiple-choice items, one of which was combined with a related constructed-response item during scaling to form a cluster. Of all of the blocks, eight contained one or more constructed-response items scored on a 0-3 scale. One item was scored on a 0-4 scale. All constructed-response items were scored by specially trained readers, as described in Chapter 5.

The design of the 1996 State Assessment in science required that each student be administered one of the 37 booklets in the design. Within each administration site, all booklets were "spiraled" together in a random sequence and distributed to students sequentially, in the order of the students' names on the Student Listing Form (see Chapter 4). As a result of the spiraling of booklets, a considerable degree of balance was achieved in the data collection process. Each block of items (and, therefore, each item) was administered to randomly equivalent samples of students within each jurisdiction.

As described in Chapter 4, a randomly selected portion of the administration sessions within each jurisdiction were observed by Westat-trained quality control monitors. Thus, within and across jurisdictions, randomly equivalent samples of students received each block of items under monitored and unmonitored administration conditions. For most jurisdictions the monitored rate was about 25 percent of the schools. For jurisdictions that were new to the state assessments (Alaska, DDESS, Nevada, Vermont, and Washington) 50 percent of the sessions were monitored. A comparison of the item statistics under monitored and unmonitored conditions is made in the tables included in the next section. Other results of monitoring are described in Chapter 4.

9.3 ITEM ANALYSES

9.3.1 Conventional Item and Test Analyses

Tables 9-3 and 9-4 contain summary statistics for each of the 11 paper-and-pencil blocks of items. The tables provide the item statistics separately for public- and nonpublic-school students. Block-level statistics are provided both overall and by serial position (first or second) of the block within booklet. To produce these tables, data from all 47 jurisdictions were aggregated and statistics were calculated using rescaled versions of the final reporting weights provided by Westat. The rescaling, carried out within each jurisdiction, constrained the sum of the sampling weights for public-school students within that jurisdiction to be equal to 2,000. The same rescaling factor was then applied to the weights of the nonpublic-school students. Use of the rescaled weights does nothing to alter the value of statistics calculated separately within each jurisdiction. However, for statistics obtained from samples that combine students from different jurisdictions, use of the rescaled weights results in a roughly equal contribution of each jurisdiction's data to the final value of the estimate. As discussed in Mazzeo (1991), equal contribution of each jurisdiction's data to the results of the IRT scaling was viewed as a desirable outcome and, as described in the scaling section below, similarly rescaled "senate" weights were used in carrying out that scaling. Hence, the item analysis statistics shown in Tables 9-3 and 9-4 are consistent with the weighting used in scaling. Section 9.8 contains more detailed information about the weights used in the 1996 State Assessment in science.

Specifically, Tables 9-3 and 9-4 show the number of students assigned each block of items, the average item score, the average biserial correlation, and the proportion of students attempting the last item in that block. The average item score for the block is the average, over items, of the score means for each of the individual items in the block. For dichotomously-scored multiple-choice and constructed-response items (0,1 scoring), these score means correspond to the proportion of students who correctly answered each item. For the cluster items and the more than 2-category constructed-response items, the score means were calculated as item score mean divided by the maximum number of points possible.

In NAEP analyses (both conventional and IRT-based), a distinction is made between missing responses at the end of each block (i.e., missing responses subsequent to the last item the student answered) and missing responses prior to the last observed response. Missing responses before the last observed response are considered intentional omissions. Intentional omissions were considered "omitted" and were treated as incorrect responses. In calculating the average score for each item, only students classified as having been presented the item were included in the denominator of the statistic. Missing responses at the end of the block are considered "not-reached," and treated as if they had not been presented to the student. The proportion of students attempting the last item of a block (or, equivalently, 1 minus the proportion of students not reaching the last item) is often used as an index of the degree of speededness associated with the administration of that block of items.

Dichotomously-scored items were analyzed using standard procedures that result in a report for each item that includes:

- for each option of the item, for examinees omitting and not reaching the item, and for the total sample of examinees:

  ⇒ number of examinees,
  ⇒ weighted percentage of examinees,
  ⇒ mean of number-correct scores, and
  ⇒ standard deviation of number-correct scores;

- p+, the proportion of examinees that received a correct score on the item (ratio of number correct to number correct plus wrong plus omitted);
- Δ, the inverse-normally transformed (1 - p+) scaled to mean 13 and standard deviation 4;
- the biserial correlation coefficient between the item and the number-correct scores; and
- the point-biserial correlation coefficient between the item and the number-correct scores.

Polytomously-scored items were analyzed using the following procedures:

- in place of p+, the ratio of the mean item score to the maximum possible item score was used;
- in place of Δ, the ratio of the mean item score to the maximum possible item score underwent the same transformation as that used on p+ to get Δ for dichotomously scored items;
- the polyserial correlation coefficient was used in place of the biserial; and
- the product-moment correlation coefficient was used in place of the point-biserial.

The average polyserial correlation is the average, over items, of the item-level polyserial correlations (r-biserial for dichotomous items) between the item and the number correct block score. For each item-level r-polyserial, total block number-correct score (including the item in question, and with students receiving zero points for all not-reached items) was used as the criterion variable for the correlation. The number correct score was the sum of the item scores where correct dichotomous items are assigned 1 and correct polytomous (or multiple-category) items are assigned the score category assigned to the student. Data from students classified as not reaching the item were omitted from the calculation of the statistic.

As is evident from Tables 9-3 and 9-4, the difficulty and the average item-to-total correlations of the blocks varied somewhat. Such variability was expected since these blocks were not created to be parallel in either difficulty or content. Based on the proportion of students attempting the last item for public-school students, no block seemed to be speeded, by the criterion of a proportion less than .8 attempting the last item. The most speeded block showed 84 percent of the students reaching the last item in the block.

These tables also indicate that there was little variability in average item scores or average polyserial correlations for each block by serial position within the assessment booklet. This suggests that serial position within booklet had a small effect on the overall difficulty of the block, with the weighted average item score being about one percent lower in the second position than in the first for public schools, and slightly more variable for the smaller sample of nonpublic-school students. As shown in Tables 9-3 and 9-4, a bare majority of the blocks showed decreases in the proportion attempting the last item when moving from the first to second position for public-school students, but only two of the 11 blocks for nonpublic-school students showed this pattern.

As mentioned earlier, to maintain rigorous standardized administration procedures across the jurisdictions, a randomly selected 25 percent of all sessions within each trend jurisdiction (50% for new jurisdictions) were observed by a Westat-trained quality control monitor. The monitoring was done in similar proportions for the public and nonpublic schools. Observations from this random portion of the sessions provided information about the quality of administration procedures and the frequency of departures from standardized procedures in the monitored sessions (see the last section in Chapter 4 for a discussion of the results of these observations). Unexpectedly large differences in results from monitored and unmonitored sessions (i.e., differences larger than those to be expected due to sampling fluctuation) could be used to identify instances of cheating, breaches of test security, or other breaks in standardization occurring in the unmonitored sessions that might threaten the validity of assessment results.

Table 9-5 gives descriptive statistics for the four hands-on task blocks, one of which was administered to each student. Block S5, which dealt with testing water-holding and pH of soils, was slightly more difficult than the first two blocks, but far fewer students attempted the final item. Block S6, which dealt with chromatography, was much more difficult than the other blocks, but this was not reflected in a large proportion of students failing to reach the last item.

When results were aggregated over all participating jurisdictions, there was little difference between the performance of students who attended monitored or unmonitored sessions. Tables 9-6 and 9-7 provide, for each block of items, the average item score, average polyserial, and the proportion of students attempting the last item for students whose sessions were monitored and students whose sessions were unmonitored, in public and nonpublic schools. Little or no difference in average item performance by session type was evident (0 to -.01 for public schools, -.03 to .02 for nonpublic schools). The difference in weighted proportions of students attempting the last item in a block ranged from .01 to -.02 for public-school students, and from .05 to -.03 for the smaller sample of nonpublic-school students.

Table 9-3
Descriptive Statistics for Each Block of Paper-and-Pencil Items by Position Within Test Booklet and Overall
Public-School Sessions, Grade 8

Statistic	Position	S7	S8	S9	S10	S11	S12	S13	S14	S15	S20	S21
Unweighted sample size												
	1	10519	10745	10537	8036	8026	8020	8046	10644	7998	7993	7968
	2	10594	10663	10686	7896	7915	7925	7943	10749	8041	7996	7979
	ALL	21113	21408	21223	15932	15941	15945	15989	21393	16039	15989	15947
Weighted average item score												
	1	0.45	0.56	0.49	0.35	0.43	0.34	0.40	0.43	0.37	0.44	0.44
	2	0.44	0.55	0.48	0.34	0.42	0.33	0.39	0.42	0.37	0.43	0.43
	ALL	0.44	0.56	0.49	0.35	0.43	0.33	0.40	0.43	0.37	0.44	0.44
Weighted average r-polyserial												
	1	0.66	0.68	0.56	0.48	0.52	0.53	0.49	0.55	0.55	0.51	0.51
	2	0.67	0.68	0.57	0.49	0.53	0.54	0.48	0.55	0.56	0.51	0.52
	ALL	0.66	0.68	0.57	0.48	0.53	0.53	0.49	0.55	0.56	0.51	0.52
Weighted proportion of students attempting last item												
	1	0.92	0.97	0.96	0.93	0.85	0.83	0.94	0.98	0.89	0.86	0.90
	2	0.90	0.96	0.94	0.92	0.86	0.84	0.92	0.98	0.88	0.86	0.92
	ALL	0.91	0.96	0.95	0.92	0.86	0.84	0.93	0.98	0.89	0.86	0.91

Table 9-4
Descriptive Statistics for Each Block of Paper-and-Pencil Items by Position Within Test Booklet and Overall
Nonpublic-School Sessions, Grade 8

Statistic	Position	S7	S8	S9	S10	S11	S12	S13	S14	S15	S20	S21
Unweighted sample size												
	1	828	809	798	632	617	620	621	814	628	584	593
	2	803	829	809	610	611	613	622	825	605	597	617
	ALL	1631	1638	1607	1242	1228	1233	1243	1639	1233	1181	1210
Weighted average item score												
	1	0.54	0.62	0.55	0.38	0.48	0.40	0.44	0.47	0.41	0.48	0.49
	2	0.53	0.60	0.52	0.38	0.49	0.38	0.44	0.49	0.40	0.47	0.47
	ALL	0.53	0.61	0.54	0.38	0.48	0.39	0.44	0.48	0.41	0.47	0.48
Weighted average r-polyserial												
	1	0.61	0.58	0.55	0.50	0.49	0.50	0.47	0.53	0.54	0.47	0.51
	2	0.63	0.65	0.55	0.50	0.50	0.51	0.47	0.51	0.52	0.50	0.49
	ALL	0.62	0.62	0.55	0.50	0.50	0.50	0.47	0.52	0.53	0.49.	0.50
Weighted proportion of students attempting last item												
	1	0.94	0.99	0.98	0.95	0.90	0.87	0.97	0.98	0.92	0.89	0.93
	2	0.95	0.98	0.98	0.97	0.91	0.88	0.97	0.98	0.92	0.88	0.95
	ALL	0.94	0.98	0.98	0.96	0.90	0.88	0.97	0.98	0.92	0.89	0.94

Table 9-5

Descriptive Statistics for Each Block of Hands-On Tasks Overall Combined Public and Nonpublic School Sessions, Grade 8

Statistic	S3	S4	S5	S6
Unweighted sample size	28947	26452	26080	20464
Weighted average item score	0.54	0.52	0.45	0.25
Weighted average r-polyserial	0.72	0.64	0.78	0.77
Weighted proportion of students attempting last item	0.76	0.83	0.43	0.80

Table 9-6
Block-Level Descriptive Statistics for Monitored and Unmonitored Public-School Sessions
Grade 8

Statistic	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S20	S21
Unweighted sample size															
Unmonitored	19355	17778	17461	17745	15612	15917	15774	11818	11784	11792	11881	15860	11934	11883	11887
Monitored	6772	6107	6076	6157	5501	5491	5449	4114	4157	4153	4108	5533	4105	4106	4060
Weighted average item score															
Unmonitored	0.53	0.52	0.44	0.25	0.44	0.55	0.48	0.34	0.42	0.33	0.40	0.42	0.37	0.44	0.43
Monitored	0.54	0.53	0.45	0.25	0.45	0.56	0.49	0.35	0.43	0.34	0.40	0.43	0.38	0.44	0.44
Weighted average r-polyserial															
Unmonitored	0.72	0.65	0.78	0.78	0.66	0.68	0.57	0.48	0.53	0.53	0.49	0.55	0.56	0.51	0.52
Monitored	0.72	0.64	0.79	0.76	0.66	0.67	0.57	0.48	0.52	0.54	0.49	0.54	0.55	0.51	0.51
Weighted proportion of students attempting last item															
Unmonitored	0.75	0.82	0.43	0.79	0.91	0.96	0.95	0.92	0.86	0.84	0.93	0.98	0.88	0.86	0.91
Monitored	0.77	0.84	0.43	0.79	0.91	0.97	0.95	0.93	0.85	0.84	0.93	0.98	0.89	0.87	0.91

Table 9-7
Block-Level Descriptive Statistics for Monitored and Unmonitored Nonpublic-School Sessions
Grade 8

Statistic	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S20	S21
Unweighted sample size															
Unmonitored	1503	1363	1322	1346	1202	1207	1181	925	907	922	917	1213	905	865	876
Monitored	524	475	478	496	429	431	426	317	321	311	326	426	328	316	334
Weighted average item score															
Unmonitored	0.61	0.57	0.52	0.28	0.53	0.61	0.54	0.38	0.49	0.39	0.43	0.48	0.41	0.48	0.47
Monitored	0.59	0.55	0.52	0.28	0.53	0.61	0.54	0.37	0.48	0.39	0.46	0.48	0.41	0.47	0.49
Weighted average r-polyserial															
Unmonitored	0.72	0.60	0.76	0.74	0.61	0.62	0.55	0.51	0.50	0.51	0.48	0.52	0.53	0.48	0.50
Monitored	0.72	0.64	0.78	0.73	0.64	0.61	0.56	0.48	0.50	0.49	0.45	0.53	0.54	0.51	0.52
Weighted proportion of students attempting last item															
Unmonitored	0.80	0.88	0.49	0.87	0.95	0.98	0.98	0.95	0.90	0.89	0.96	0.98	0.92	0.89	0.93
Monitored	0.81	0.89	0.48	0.88	0.92	0.99	0.99	0.97	0.90	0.84	0.99	0.97	0.91	0.90	0.95

Table 9-8 summarizes the differences between monitored and unmonitored average item scores for the jurisdictions. These are mean differences within a jurisdiction averaged over all items in all the blocks. The information in this table combines public and nonpublic school data. The median difference was essentially zero. For 16 jurisdictions, the differences were negative (i.e., students from unmonitored sessions scored higher than students from monitored sessions). None were larger in absolute magnitude than .01. The results indicate that across jurisdictions, there were only small differences between monitored and unmonitored sessions.

Table 9-8
The Effect of Monitoring Sessions by Jurisdiction:
Average Jurisdiction Item Scores for Monitored and Unmonitored Sessions, Grade 8

	Unmonitored	**Monitored**	**Monitored - Unmonitored**
Alabama	0.081	0.080	-0.001
Alaska	0.091	0.092	0.001
Arizona	0.083	0.087	0.004
Arkansas	0.085	0.082	-0.003
California	0.081	0.080	-0.001
Colorado	0.092	0.094	0.002
Connecticut	0.094	0.094	0.000
Delaware	0.086	0.082	-0.004
DoDEA/DDESS	0.090	0.091	0.001
DoDEA/DoDDS	0.092	0.093	0.001
District of Columbia	0.065	0.072	0.007
Florida	0.080	0.084	0.004
Georgia	0.081	0.085	0.004
Guam	0.069	0.077	0.008
Hawaii	0.076	0.079	0.003
Indiana	0.091	0.090	-0.001
Iowa	0.095	0.097	0.002
Kentucky	0.085	0.089	0.004
Louisiana	0.076	0.078	0.002
Maine	0.100	0.097	-0.003
Maryland	0.087	0.088	0.001
Massachusetts	0.093	0.099	0.006
Michigan	0.092	0.091	-0.001
Minnesota	0.096	0.096	0.000
Mississippi	0.074	0.074	0.000
Missouri	0.092	0.089	-0.003
Montana	0.097	0.099	0.002
Nebraska	0.096	0.096	0.000
Nevada	0.084	0.085	0.001
New Hampshire	0.098	0.098	0.000
New Jersey	0.091	0.093	0.002
New Mexico	0.081	0.083	0.002
New York	0.086	0.085	-0.001
North Carolina	0.085	0.084	-0.001
North Dakota	0.101	0.096	-0.005
Oregon	0.091	0.100	0.009

Table 9-8 (continued)

The Effect of Monitoring Sessions by Jurisdiction:
Average Jurisdiction Item Scores for Monitored and Unmonitored Sessions, Grade 8

	Unmonitored	Monitored	Monitored - Unmonitored
Rhode Island	0.089	0.089	0.000
South Carolina	0.080	0.077	-0.003
Tennessee	0.083	0.085	0.002
Texas	0.086	0.086	0.000
Utah	0.094	0.092	-0.002
Vermont	0.095	0.094	-0.001
Virginia	0.088	0.085	-0.003
Washington	0.089	0.086	-0.003
West Virginia	0.084	0.088	0.004
Wisconsin	0.098	0.101	0.003
Wyoming	0.094	0.097	0.003
		mean:	0.001
		median:	0.001
		minimum:	-0.005
		maximum:	0.009
		1st Quartile:	-0.001
		3rd Quartile:	0.003

As has been the case since the 1994 Trial State Assessment in reading, the 1996 State Assessment in science included students sampled from nonpublic schools. The nonpublic-school population that was sampled included students from Catholic schools, private religious schools and private non-religious schools (all referred to by the term 'nonpublic schools'). Table 9-9 shows the difference between public and nonpublic schools with respect to average item statistics, average r-polyserial correlation, and average proportion of students attempting the last item in a block. As with the monitored/unmonitored comparisons, results were aggregated over all participating jurisdictions.

All of the 47 jurisdictions that participated in the State Assessment in science had public-school samples, while 33 of the 47 jurisdictions had nonpublic-school samples that met reporting requirements.

Consistent differences are evident between the public- and nonpublic-school groups. Table 9-9 indicates that the difference in average item score between public- and nonpublic-school students (i.e., public block mean minus nonpublic block mean) ranged from -.03 to -.09 with an average of -.05, indicating that public-school students were generally lower in average item score. The public/nonpublic difference in average item-to-total block correlation (the average r-polyserial) ranged from -.02 to .06 with an average of .02, indicating that public-school students generally had a somewhat higher item-to-total correlation. As for the proportion of students attempting the last item, public minus nonpublic differences ranged from 0 to -.08 with an average of -.04, indicating that somewhat fewer students in public schools attempted the last item.

Table 9-9

Block-Level Descriptive Statistics for Overall Public- and Nonpublic-School Sessions

Grade 8

Statistic	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S20	S21
Unweighted sample size															
Public	26127	23885	23537	23902	21113	21408	21223	15932	15941	15945	15989	21393	16039	21393	16039
Nonpublic	2027	1838	1800	1842	1631	1638	1607	1242	1228	1233	1243	1639	1233	1639	1233
Weighted average item score															
Public	0.53	0.52	0.44	0.25	0.44	0.56	0.49	0.35	0.43	0.33	0.40	0.43	0.37	0.43	0.37
Nonpublic	0.61	0.56	0.52	0.28	0.53	0.61	0.54	0.38	0.48	0.39	0.44	0.48	0.41	0.48	0.41
Weighted average r-polyserial															
Public	0.72	0.65	0.78	0.77	0.66	0.68	0.57	0.48	0.53	0.53	0.49	0.55	0.56	0.55	0.56
Nonpublic	0.72	0.61	0.76	0.74	0.62	0.62	0.55	0.50	0.50	0.50	0.47	0.52	0.53	0.52	0.53
Weighted Proportion of students attempting last item															
Public	0.75	0.83	0.43	0.79	0.91	0.96	0.95	0.92	0.86	0.84	0.93	0.98	0.89	0.98	0.89
Nonpublic	0.80	0.89	0.48	0.87	0.94	0.98	0.98	0.96	0.90	0.88	0.97	0.98	0.92	0.98	0.92

Tables 9-10 and 9-11 summarize the distribution over jurisdictions of the mean item scores for public- and nonpublic-school students. Across jurisdictions, the average public-school scores were quite varied with the difference between the first and third quartiles (an indication of how variable scores are) as high as .059 (for *earth science*). The spread of scores for nonpublic-school students was somewhat smaller. The differences between public- and nonpublic-school scores was comparable to the difference between jurisdictions, ranging from .047 (for *life science*) to .065 (for *earth science*). Table 9-12 summarizes the distribution for the two school types combined.

Table 9-10
Distribution of Jurisdiction Mean Item Scores by Fields of Science Scale
Public Schools

	Earth Science	Physical Science	Life Science
Mean:	0.386	0.416	0.356
Median:	0.387	0.420	0.360
1st Quartile:	0.362	0.395	0.337
3rd Quartile:	0.421	0.452	0.381
Minimum:	0.245	0.282	0.246
Maximum:	0.452	0.485	0.404
Number of jurisdictions:	47	47	47

Table 9-11
Distribution of Jurisdiction Mean Item Scores by Fields of Science Scale
Nonpublic Schools

	Earth Science	Physical Science	Life Science
Mean:	0.449	0.473	0.405
Median:	0.452	0.472	0.407
1st Quartile:	0.422	0.452	0.390
3rd Quartile:	0.478	0.493	0.420
Minimum:	0.367	0.407	0.347
Maximum:	0.512	0.531	0.468
Number of jurisdictions:	33	33	33

9.3.2 Differential Item Functioning (DIF) Analyses

Before the state data were analyzed further, differential item functioning (DIF) analyses were carried out on the 1996 NAEP science data from the national samples at grades 4, 8, and 12. The purpose of these analyses was to identify items that were differentially difficult for members of various subgroups with comparable overall scores and to reexamine such items with respect to their fairness and their appropriateness for inclusion in the scaling process. A separate DIF analysis was not conducted on the state data since the results of the national DIF analysis were assumed to hold for the state sample. The information in this section provides a brief description of the national DIF analysis, though a more thorough discussion and presentation of results based on the national assessment will appear in the forthcoming technical report for that assessment.

The DIF analyses of the dichotomous items were based on the Mantel-Haenszel chi-square procedure, as adapted by Holland and Thayer (1988). The procedure tests the statistical hypothesis that the odds of correctly answering an item are the same for two groups of examinees that have been matched on some measure of proficiency (usually referred to as the matching criterion). The DIF analyses of the polytomous items were based on the Mantel (1963) procedure and the Somes (1986) chi-square test. These procedures compare proportions of matched examinees from each group in each polytomous item response category. The groups being compared are often referred to as the focal group (usually a minority or other group of interest, such as Black examinees or female examinees) and the reference group (usually White examinees or male examinees).

For each dichotomous item in the assessment, an estimate of the Mantel-Haenszel common odds-ratio, expressed on the ETS delta scale for item difficulty was produced. The estimates indicate the difference between reference group and focal group item difficulties (measured in ETS delta scale units), and typically run between about +3 and -3. Positive values indicate items that are differentially easier for the focal group than the reference group after making an adjustment for the overall mean item score for the two groups. Similarly, negative values indicate items that are differentially harder for the focal group than the reference group. It is common practice at ETS to categorize each item into one of three categories (Petersen, 1988): “A” (items exhibiting no DIF), “B” (items exhibiting a weak indication of DIF), or “C” (items exhibiting a strong indication of DIF). Items in category A have Mantel-Haenszel common odds ratios on the delta scale that do not differ significantly from 0 at the alpha = .05 level or are less than 1.0 in absolute value. Category C items are those with Mantel-Haenszel values that are significantly greater than 1 and larger than 1.5 in absolute magnitude. Other items are categorized as B items. A plus sign (+) indicates that items are differentially easier for the focal group; a minus sign (-) indicates that items are differentially more difficult for the focal group.

The ETS/NAEP DIF procedure for polytomous items incorporates the Mantel-Haenszel ordinal procedure. Polytomous items are categorized as “AA,” “BB,” and “CC,” generalizations of the dichotomous A, B, and C categories.

Following standard practice at ETS for DIF analyses conducted on final test forms, all C and CC items were reviewed by a committee of trained test developers and subject-matter specialists. Such committees are charged with making judgments about whether or not the differential difficulty of an item is *unfairly* related to group membership. As pointed out by Zieky (1993):

> It is important to realize that *DIF* is not a synonym for *bias*. The item response theory based methods, as well as the Mantel-Haenszel and standardization methods of DIF detection, will identify questions that are not measuring the same dimension(s) as the bulk of the items in the matching criterion....Therefore, judgement is required to determine whether or not the difference in difficulty shown by a DIF index is *unfairly* related to group membership. The judgement of fairness is based on whether or not the difference in difficulty is believed to be related to the construct being measured....The fairness of an item depends directly on the purpose for which a test is being used. For example, a science item that is differentially difficult for women may be judged to be fair in a test designed for certification of science teachers because the item measures a topic that every entry-level science teacher should know. However, that same item, with the same DIF value, may be judged to be unfair in a test of general knowledge designed for all entry-level teachers.
> (p. 340)

A committee assembled to review NAEP items that were identified as C or CC items. The committee included both ETS staff and outside members with expertise in the field. It was the committee's judgment that none of the items was functioning differentially for the national grade 8 data due to factors irrelevant to test objectives.

9.4 ITEM RESPONSE THEORY (IRT) SCALING

Items were sorted into three distinct sets, one for each of the three science fields of science scales. Table 9-12 describes the mean item scores for the items comprising each of the three fields of science scales. In contrast to previous item mean tables, this table describes the item means averaged over jurisdictions for each item, rather than item means averaged over items for each jurisdiction. The averages are based on the entire sample of students in the State Assessment and use a weighting in which monitored and unmonitored sessions within jurisdictions were given equal weights. The median values indicate that students found the set of *earth science* items to be the most difficult.

Separate IRT-based scales corresponding to each of the item sets defined above were developed using the scaling models described in Chapter 8. Three scales were produced by separately calibrating the sets of items classified in each of the three fields of science scales.

For the reasons discussed in Mazzeo (1991), for each scale, parameters for each item were estimated for the entire data set and used for all jurisdictions. Item parameter estimation was carried out using a 25 percent random sample of the public-school students participating in the 1996 State Assessment. It included equal numbers of students from each participating jurisdiction, with 25 percent taken from the monitored sessions and 75 percent taken from the unmonitored sessions. The sample consisted of 24,750 students with 550 students being sampled from each of the 45 jurisdictions (excluding DDESS and DoDDS schools) included in the sample.[4] All calibrations were carried out using the rescaled sampling weights described in Section 9.3.1 in an effort to ensure that each jurisdiction's data contributed equally to the determination of the item parameter estimates.

To the extent that items may have functioned differently in monitored and unmonitored sessions, the single set of item parameters obtained define a set of item characteristic curves "averaged over" the two types of sessions. Tables 9-3 and 9-4 (shown earlier) presented block-level item statistics that suggested little, if any, differences in item functioning by session type. The item calibration was carried out only with the public-school students sampled from each jurisdiction. In order to gauge whether items functioned differentially between public and nonpublic sessions, a DIF analysis was done with public-school students defined as the reference group and nonpublic-school students defined as the focal group. The procedure was the same as that for testing DIF for minority and gender groups in the national sample (as was described in Section 9.3 above). Not a single item was identified as showing significant DIF in this analysis (i.e., there were no C nor CC items). As a result, it was concluded that items did not give evidence of functioning differentially for public- and nonpublic-school students so item parameters based on the public-school sample were used to set the scales.

[4] Students from DDESS and DoDDS jurisdictions were not included in the 25% subsample for scaling.

Table 9-12
Distribution of Item Mean Scores Averaged Across All Students in the State Assessment

	Earth Science	Physical Science	Life Science
Mean:	0.418	0.417	0.404
Median:	0.406	0.415	0.343
1st Quartile:	0.230	0.273	0.207
3rd Quartile:	0.551	0.557	0.574
Minimum:	0.034	0.009	0.034
Maximum:	0.919	0.943	0.943
Number of items:	65	62	65

9.4.1 Item Parameter Estimation

For each field of science scale, item parameter estimates were obtained from the NAEP BILOG/PARSCALE program, which combines Mislevy and Bock's (1982) BILOG and Muraki and Bock's (1991) PARSCALE computer programs. The program uses marginal estimation procedures to estimate the parameters of the one-, two-, and three-parameter logistic models, and the generalized partial credit model described by Muraki (1990) (see Chapter 8).

All multiple-choice items were dichotomously scored (scored 0,1) and were scaled using the three-parameter logistic model. Omitted responses to multiple-choice items were treated as fractionally correct, with the fraction being set to the reciprocal of the number of response options for an item. All constructed-response items with two categories were dichotomously scored and were scaled using the two-parameter logistic model with the lower asymptote parameter set at 0. Omitted responses to these items were treated as incorrect.

A key assumption associated with IRT scales is that of conditional independence. Conditional on proficiency, examinee's item responses are assumed to be independent. When sets of items are logically dependent on each other, or are based on a single stimulus, this assumption can be violated to a degree that results in aberrant scaling results. Because of the dependency of related items, two cluster items were scored directly as single polytomous items. In order to avoid possible problems with other inter-item dependencies, two additional cluster items were created by combining examinee responses to sets of related items into a single score for each set. The cluster items, rather than their original constituent items, were used in scaling the 1996 science assessment. In defining these additional cluster items, examinees were categorized by a sum of the "credit" they received for the two contributing items: examinees omitting all constituents of the cluster item were placed in the "omit" category of the cluster item, and examinees classified as "not reaching" all constituent parts were placed in the "not reached" category. All cluster items were scaled using the generalized partial credit model.

There was a total of 96 multi-category constructed response items other than the cluster items just described. Each of these items was also scaled using the generalized partial credit model. These items had two, three, four, or five categories of partial credit. Scoring levels for polytomously scored items were labeled and categorized as shown in Table 9-13. For all items included in scaling, the percentage of public-school students omitting a response ranges from 0.00 to 27.10, and the percentage

Table 9-13
Scoring Levels

Score	3-Category	4-Category	5-Category	6-Category
5	—	—	—	Complete
4	—	—	Complete	Essential
3	—	Complete	Essential	Adequate
2	Complete	Partially correct	Partially Correct	Partially Correct
1	Partially Correct	Unsatisfactory	Unsatisfactory	Unsatisfactory
0	Wrong, off-task, or omitted	Wrong, off-task, or omitted	Wrong, off-task, or omitted	Wrong, off-task, or omitted

of public-school students giving off-task responses ranges from 0.00 to 5.09. The percentage of students attempting the last item in each block, and, by subtraction, the percentage of students not reaching the last item in each block is presented in Table 9-9.

Empirical Bayes modal estimates of all item parameters were obtained from the BILOG/PARSCALE program. Prior distributions were imposed on item parameters with the following starting values: thresholds (normal [0,2]); slopes (log-normal [0,.5]); and asymptotes (two-parameter beta with parameter values determined as functions of the number of response options for an item and a weight factor of 50). The locations (but not the dispersions) were updated at each program estimation cycle in accordance with provisional estimates of the item parameters.

Item parameter estimation proceeded in two phases. First, the subject ability distribution was assumed fixed (normal [0,1]) and a stable solution was obtained. The parameter estimates from this solution were then used as starting values for a subsequent set of runs in which the subject ability distribution was freed and estimated concurrently with item parameter estimates. After each estimation cycle, the subject ability distribution was restandardized to have a mean of zero and standard deviation of one. Correspondingly, parameter estimates for that cycle were also linearly restandardized.

During and subsequent to item parameter estimation, evaluations of the fit of the IRT models were carried out for each of the items in the item pool. These evaluations were conducted to determine the final composition of the item pool making up the scales by identifying misfitting items that should not be included. Evaluations of model fit were based primarily on graphical analyses. For dichotomously-scored multiple-choice and two-category response items, model fit was evaluated by examining plots of estimates of the expected conditional (on θ) probability of a correct response that do not assume a two-parameter or three-parameter logistic model versus the probability predicted by the estimated item characteristic curve (see Mislevy & Sheehan, 1987, p. 302). For the cluster items and multiple-category constructed-response items, similar plots were produced for each item category characteristic curve.

As with most procedures that involve evaluating plots of data versus model predictions, a certain degree of subjectivity is involved in determining the degree of fit necessary to justify use of the model. There are a number of reasons why evaluation of model fit relied primarily on analyses of plots rather than seemingly more objective procedures based on goodness-of-fit indices such as the “pseudo chi-squares” produced in BILOG (Mislevy & Bock, 1982). First, the exact sampling distributions of these indices when the model fits are not well understood, even for fairly long tests. Mislevy and Stocking (1987) point out that the usefulness of these indices appears particularly limited in situations like NAEP where examinees have been administered relatively short tests. A study by Stone, Mislevy, and Mazzeo

(1994) using simulated data suggests that the correct reference chi-square distributions for these indices have considerably fewer degrees of freedom than the value indicated by the BILOG/PARSCALE program and require additional adjustments of scale. However, it is not yet clear how to estimate the correct number of degrees of freedom and necessary scale factor adjustment factors. Consequently, pseudo chi-square goodness-of-fit indices are used only as rough guides in interpreting the severity of model departures.

Second, as discussed in Chapter 8, it is almost certainly the case that, for most items, item-response models hold only to a certain degree of approximation. Given the large sample sizes used in NAEP and the State Assessment, there will be sets of items for which one is almost certain to reject the hypothesis that the model fits the data (since the likelihood of rejecting the null increases with sample size) even though departures are minimal in nature or involve kinds of misfit unlikely to impact on important model-based inferences about student achievement. In practice, it is always wise to temper statistical decisions with judgments about the severity of model misfit and the potential impact of such misfit on final results.

In making decisions about excluding items from the final scales, a balance was sought between being too stringent, hence deleting too many items and possibly damaging the content representativeness of the pool of scaled items, and too lenient, hence including items with model fit poor enough to invalidate the types of inferences made from NAEP results. Items that clearly did not fit the model were not included; however, a certain degree of misfit was tolerated for a number of items included in the final scales. The scaling for the 1996 State Assessment in science began by incorporating all adjustments and deletions resulting from the 1996 national science assessment.

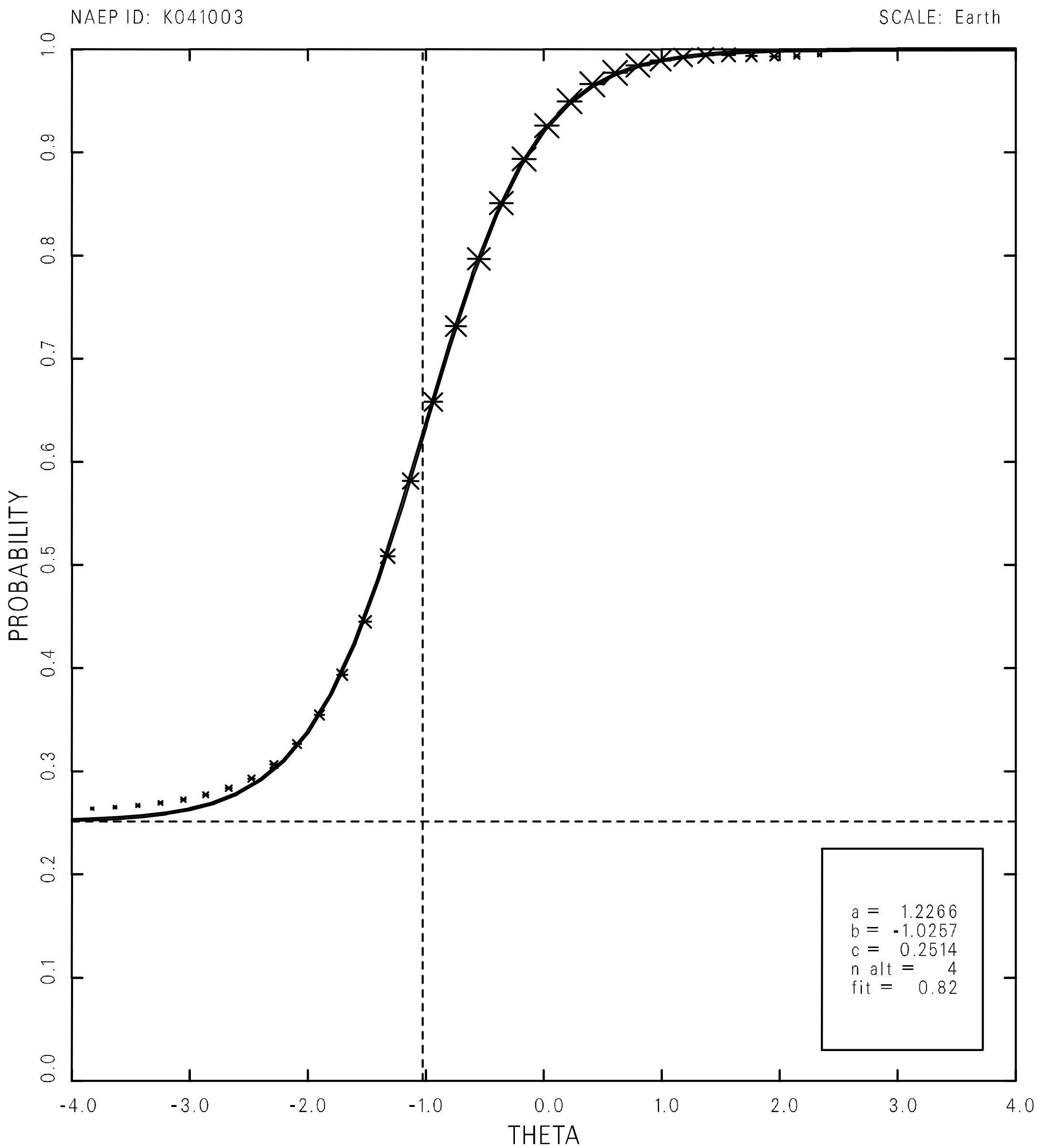
For the large majority of the items, the fit of the model was extremely good. Figure 9-1 provides a typical example of what the plots look like for a dichotomously-scored item in this class of items. The plot that is shown is for an item from the earth science scale. In the plot, the y-axis indicates the probability of a correct response and the x-axis indicates scale score level (θ) when the distribution of scale scores has a mean of zero and a standard deviation of one. The curve comprised of asterisks shows estimates of the conditional (on θ) probability of a correct response conditioned on θ. No assumptions of a logistic form is made for this curve, hence, the points on the curve are referred to subsequently as empirical or nonlogistic-based estimates. The sizes of the asterisks are proportional to the estimated density of the theta distribution at the indicated value. The solid curve shows the estimated theoretical item response function. The item response function provides estimates of the conditional probability of a correct response based on an assumed logistic form. The vertical dashed line indicates the estimated location parameter (b) for the item and the horizontal dashed line indicates the estimated lower asymptote (c). Also shown in the plot are the actual values of the item parameter estimates (lower right-hand corner). As is evident from the plot, the empirical or nonlogistic-based item trace is in extremely close agreement with the model-based item response function.

Figure 9-2 provides an example of a plot for a 4-category extended constructed-response item exhibiting good model fit. Like the plots for the dichotomously-scored items, this plot shows two estimates of each item category characteristic curve, one set that does not assume the partial credit model (the empirical trace shown as asterisks) and one that does (the theoretical trace shown as solid curves). As is the case with Figure 9-1, the two sets of estimates agree quite well, although there is a slight tendency for the nonlogistic-based estimates for category two to be somewhat higher than the model-based estimates in the region near theta equals 0.0.

As discussed above, some of the items retained for the final scales display some degree of model misfit. Figures 9-3 (a dichotomously-scored multiple-choice item) and 9-4 (an extended constructed-response item) provide typical examples of such items. The slightly nonlogistic form of the empirical curve in Figure 9-3, above the model-based function in the middle range, and below the curve outside this range, is not uncommon among the small number of items exhibiting some lack of fit. Note that in Figure 9-4, the empirical curve lies above the theoretical curve in the lower part of the ability scale for the lowest category, but below the theoretical curve in the same scale score range for the next higher category. Combining these two categories would have improved the model fit, but it was judged that the misfit was not sufficiently pronounced in this case to warrant such collapsing. In general, good agreement between empirical and theoretical item traces were found for the regions of the theta scale that include most of the examinees. Misfit was confined to conditional probabilities associated with theta values in the tails of the subject ability distributions.

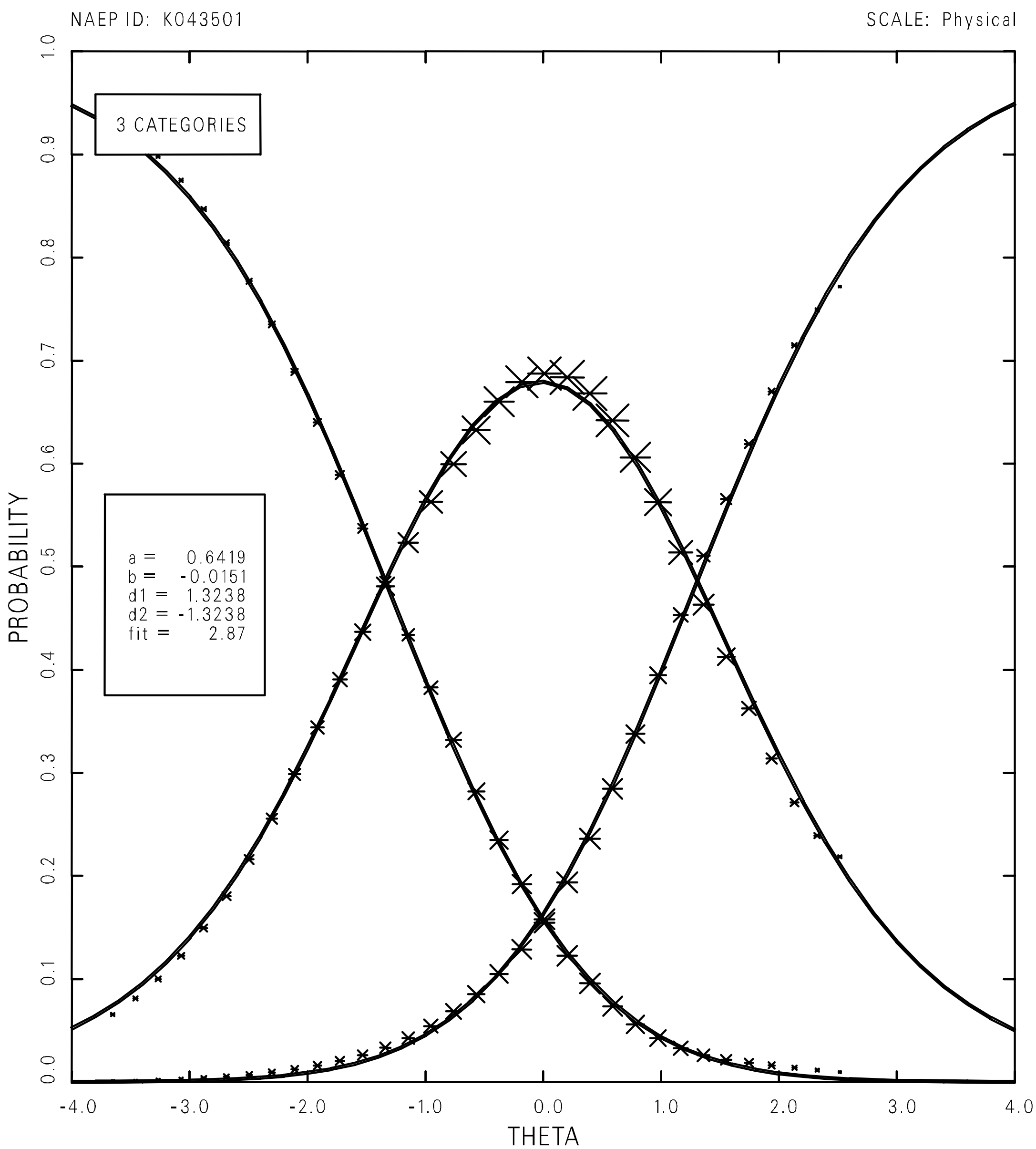
Although, similar results in regard to item fit were found for state and national data, there was one item for which the state analysis showed potential lack of item fit and the national analysis did not. Figure 9-5 contains the item plot for this item in the *earth science* scale (item K044101) after estimating the parameters with the prior unconstrained to the normal distribution. All of the items in this run were treated in the way indicated by the fit of the items to the national data. As can be seen in the plot, the middle and highest categories show a level of misfit not seen in the lowest category. Categories 2 and 3 were combined to yield a dichotomous item. The plot in Figure 9-6 shows that the fit of the newly-formed dichotomous item is very good. This was the only case in which the state analysis yielded a different result from the original national analysis. After a decision about the treatment of this item was made based on the state data, the treatment of the item in the national scale was adjusted to match the treatment in the state scale. So, all items were treated in the same way for both national and state scales.

Figure 9-1
Plot Comparing Empirical and Model-Based Estimates of Item Response Functions for Binary Scored (Multiple-Choice) Items Exhibiting Good Model Fit*



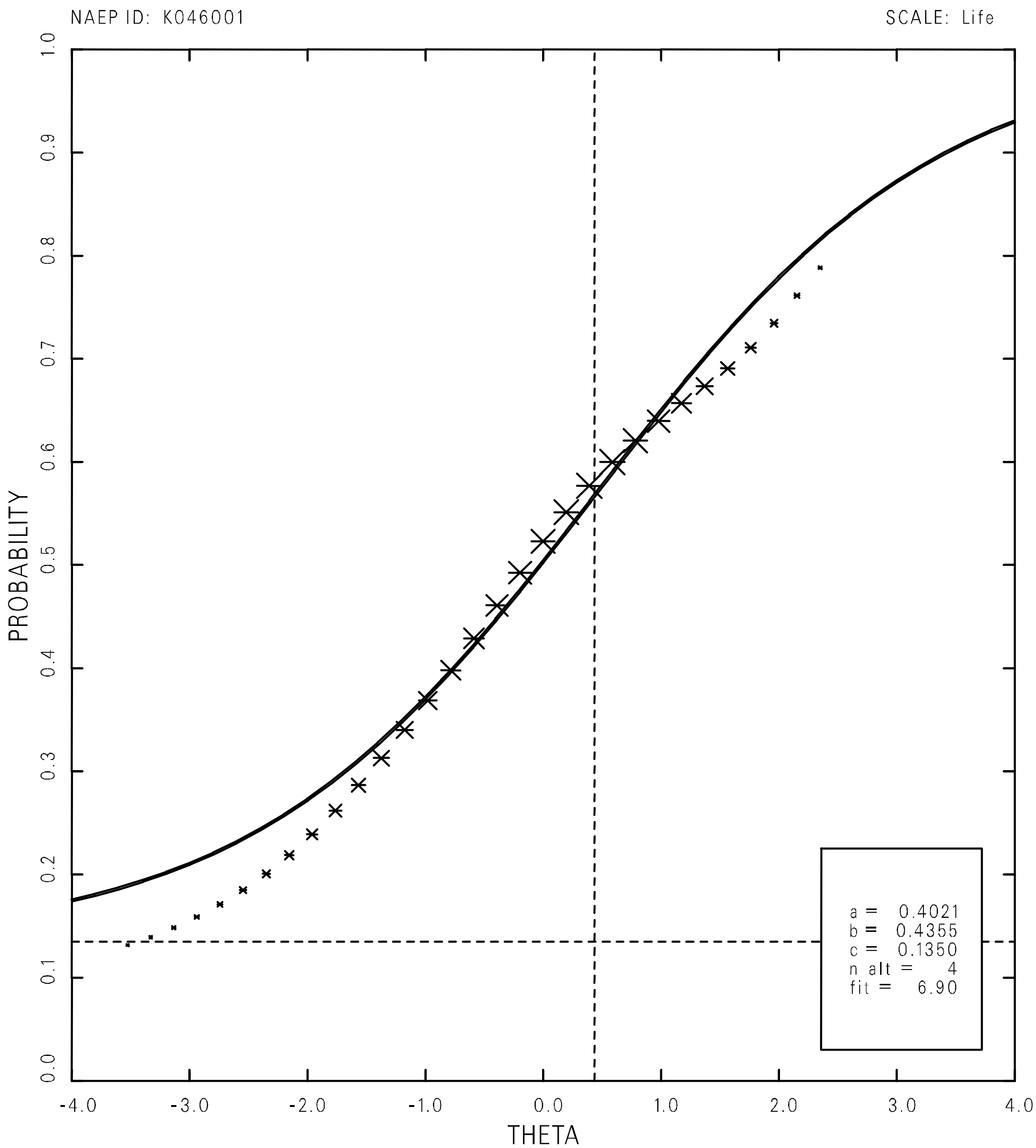
*Asterisks indicate estimated conditional probabilities obtained without assuming a logistic form; the solid curve indicates estimated item response function assuming a logistic form.

Figure 9-2
Plot Comparing Empirical and Model-Based Estimates of Item Category Characteristic Curves for a Polytomously Scored Item Exhibiting Good Model Fit*



*Asterisks indicate estimated conditional probabilities obtained without assuming a model-based form; the solid curve indicates estimated item response function assuming a model-based form. The number of categories in this figure includes the zero category contrary to usage in the text of this report.

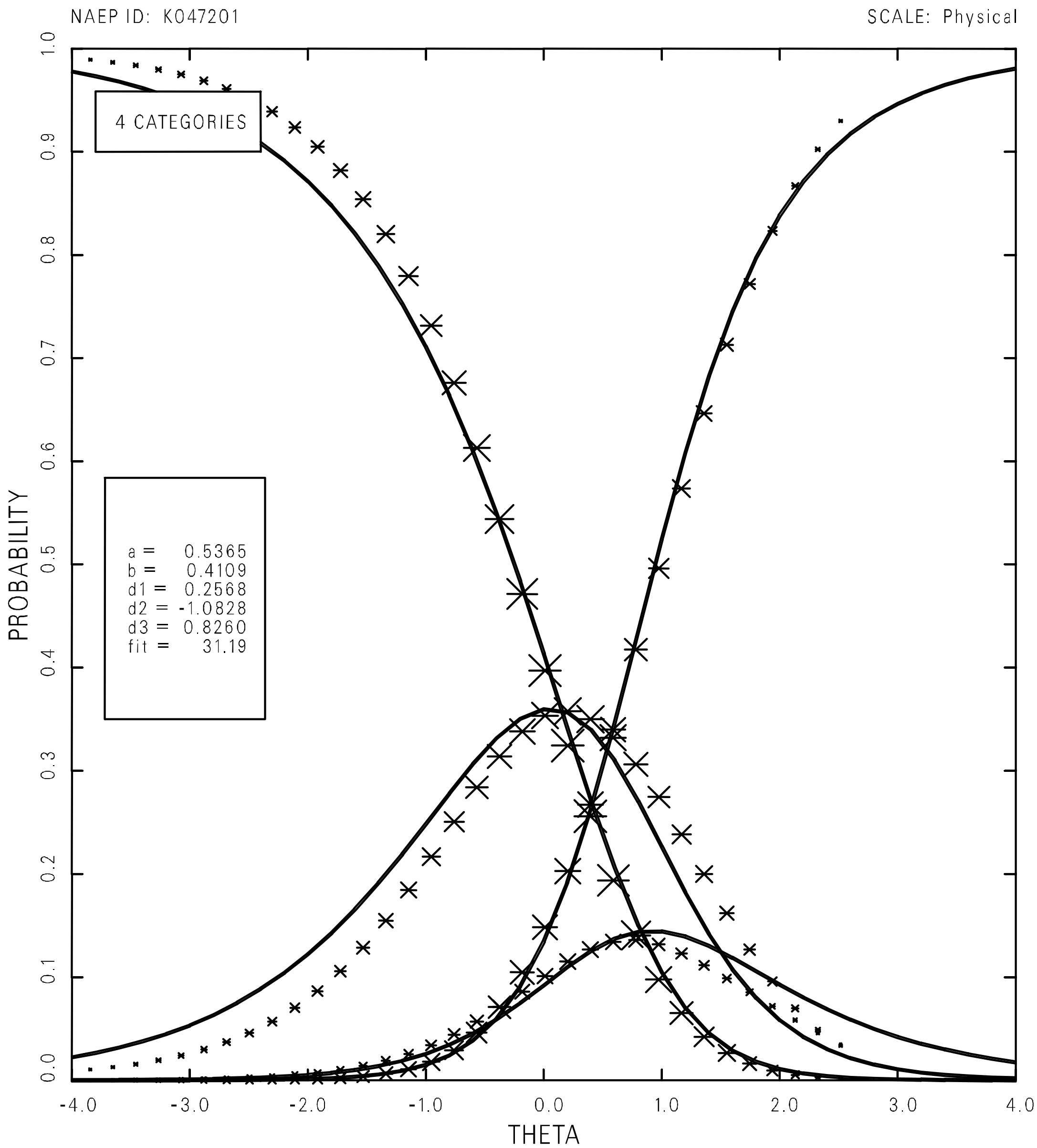
Figure 9-3
Plot Comparing Empirical and Model-Based Estimates of Item Response Functions for Binary-Scored (Multiple-Choice) Item Exhibiting Some Model Misfit*



*Asterisks indicate estimated conditional probabilities obtained without assuming a logistic form; the solid curve indicates estimated item response function assuming a logistic form.

Figure 9-4

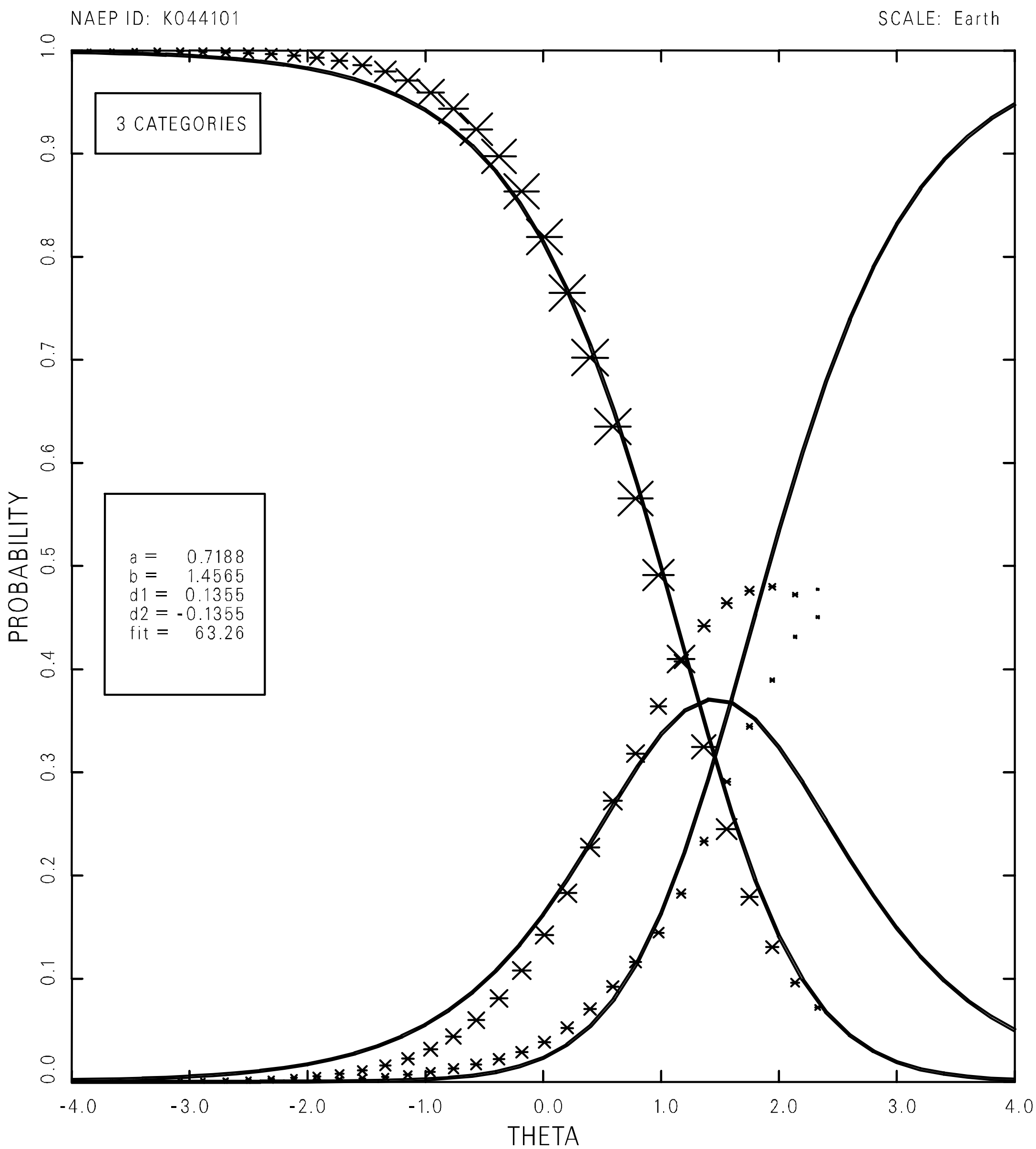
Plot Comparing Empirical and Model-Based Estimates of Item Category Characteristic Curves for a Polytomously Scored Item Exhibiting Some Model Misfit*



*Asterisks indicate estimated conditional probabilities obtained without assuming a model-based form; the solid curve indicates estimated item response function assuming a model-based form. The number of categories in this figure includes the zero category contrary to usage in the text of this report.

Figure 9-5

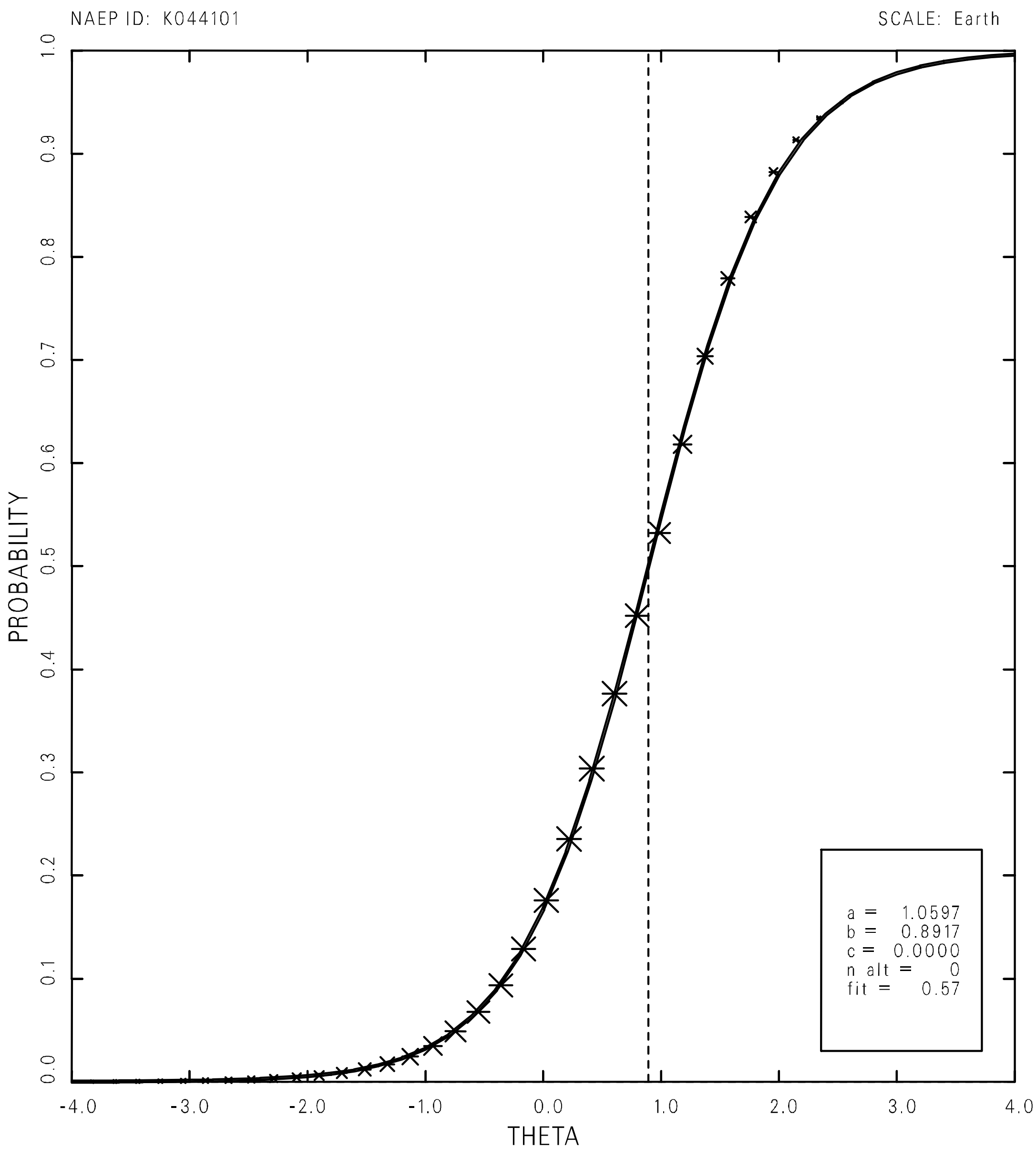
Plot Comparing Empirical and Model-Based Estimates of Item Category Characteristic Curves for a Polytomously Scored Item (K044101) Exhibiting Poor Model Fit*



*Asterisks indicate estimated conditional probabilities obtained without assuming a model-based form; the solid curve indicates estimated item response function assuming a model-based form. The number of categories in this figure includes the zero category contrary to usage in the text of this report.

Figure 9-6

Plot Comparing Empirical and Model-Based Estimates of Item Category Characteristic Curves for Polytomously Scored Item (K044101) After Collapsing Categories 2 and 3*



*Asterisks indicate estimated conditional probabilities obtained without assuming a logistic form; the solid curve indicates estimated item response function assuming a logistic form.

Table 9-14 lists the items that received special treatment during the scaling process. Included in the table are the block locations and item numbers for the items that were combined into cluster items as well as for those that were excluded from the final scales. These items received identical special treatment in the development of the 1996 national scales. No other items in either assessment received special treatment. The IRT parameters for the items included in the State Assessment are listed in Appendix D.

9.5 ESTIMATION OF STATE AND SUBGROUP PROFICIENCY DISTRIBUTIONS

The proficiency distributions in each jurisdiction (and for important subgroups within each jurisdiction) were estimated by using the multivariate plausible values methodology and the corresponding CGROUP computer program (described in Chapter 8). The 1996 State Assessment used an enhanced version of this program, based on modifications described by Thomas (1993), to estimate the proficiency distribution for each jurisdiction. As described in Chapter 8, the CGROUP program estimates scale score distributions using information from student item responses, measures of student background variables, and the item parameter estimates obtained from the BILOG/PARSCALE program.

The enhancements included in the 1992 version of CGROUP included the replacement of Monte Carlo integration by analytic calculations, new methods for computing student-level posterior means and variances, and the generation of values from student-level posterior distributions for the imputation of student-level scale score values. Simulation studies indicate that the enhanced CGROUP produces more accurate estimates of scale (i.e. fields of science scale) variances and correlations (Thomas, 1993) than did the previous versions of CGROUP.

For the reasons discussed in Mazzeo (1991), separate conditioning models were estimated at each grade for each jurisdiction. This resulted in the estimation of 47 distinct conditioning models. The student background variables included in each jurisdiction's model (denoted y in Chapter 8) were principal component scores derived from the within-state correlation matrix of selected main-effects and two-way interactions associated with a wide range of student, teacher, school, and location variables. The main-effect and interaction variables are listed in Appendix C. A set of five multivariate plausible values was drawn for each student who participated in the State Assessment.

Reporting each jurisdiction's results required analyses describing the relationships between scale scores and a large number of background variables. The background variables included student demographic characteristics (e.g., the race/ethnicity of the student, highest level of education attained by parents), students' perceptions about science, student behavior both in and out of school (e.g., amount of TV watched daily, amount of science homework done each day), the type of science class being taken (e.g., *earth science*), the amount of classroom emphasis on various topics included in the assessment provided by the students' teachers, and a variety of other aspects of the students' background and preparation, the background and preparation of their teachers, and the educational, social, and financial environment of the schools they attended.

Table 9-14
Items from the 1996 State Assessment in Science Receiving Special Attention[1]

NAEP ID	Block/Item Number	Field of Science	Action	Reason
K040607	S3 - 2A	Physical Science	Drop to Form Cluster Item	Logical Dependency
K040608	S3 - 2G	Physical Science	Drop to Form Cluster Item	Logical Dependency
K040609	S3 - 2M	Physical Science	Drop to Form Cluster Item	Logical Dependency
K040610	S3 - 2S	Physical Science	Drop to Form Cluster Item	Logical Dependency
K031301	S6 - 6	Physical Science	Drop	Dependency
K036401	S14 - 10A	Earth Science	Drop to Form Cluster Item	Logical Dependency
K036403	S14 - 10G	Earth Science	Drop to Form Cluster Item	Logical Dependency
K036404	S14 - 10M	Earth Science	Drop to Form Cluster Item	Logical Dependency
K040601	S3- 1	Physical Science	Collapse Categories 0,1,2 becomes 0,0,1	Lack of Fit
K040702	S4 - 3	Physical Science	Collapse Categories: 0,1,2,3 becomes 0,1,1,2	Lack of Fit
K040705	S4 - 4	Physical Science	Collapse Categories: 0,1,2 becomes 0,1,1	Zero Frequency
K031306	S6 - 9	Physical Science	Collapse Categories: 0,1,2 becomes 0,1,1	Zero Frequency
K043603	S11 - 16	Physical Science	Collapse Categories: 0,1,2 becomes 0,1,1	Zero Frequency
K040711	S4 - 12	Earth Science	Drop to Form Cluster Item	Dependency
K040712	S4 - 13	Earth Science	Drop to Form Cluster Item	Dependency
K040803	S5 - 2	Earth Science	Collapse Categories: 0,1,2 becomes 0,1,1	Lack of Fit
K040901	S7 - 1	Earth Science	Collapse Categories: 0,1,2 becomes 0,1,1	Lack of Fit
K040905	S7 - 5	Earth Science	Collapse Categories: 0,1,2 becomes 0,0,1	Lack of Fit
K049403	S13 - 15	Earth Science	Drop to Form Cluster Item	Dependency
K049404	S13 - 16	Earth Science	Drop to Form Cluster Item	Dependency
K044101	S20 - 5	Earth Science	Collapse Categories: 0,1,2 becomes 0,1,1	Lack of Fit
K044401	S20 - 8	Earth Science	Collapse Categories: 0,1,2 becomes 0,1,1	Lack of Fit
K041306	S8 - 6	Life Science	Collapse Categories: 0,1,2 becomes 0,1,1	Lack of Fit
K031603	S9 - 3	Life Science	Collapse Categories: 0,1,2 becomes 0,0,1	Lack of Fit
K031611	S9 - 11	Life Science	Collapse Categories: 0,1,2 becomes 0,1,1	Zero Frequency
K042602	S10 - 15	Life Science	Collapse Categories: 0,1,2,3 becomes 0,1,1,2	Lack of Fit
K049301	S13 - 12	Life Science	Collapse Categories: 0,1,2 becomes 0,1,1	Zero Frequency
K037001	S15 - 1	Life Science	Drop	Lack of Fit

[1]These items were adjusted during the scoring process, before scaling.

As described in the previous chapter, to avoid biases in reporting results and to minimize biases in secondary analyses, it is desirable to incorporate measures of a large number of independent variables in the conditioning model. When expressed in terms of contrast-coded main effects and interactions, the number of variables to be included totaled 1,044. As stated earlier, Appendix C provides a listing of the full set of contrasts defined at each grade. These contrasts were the common starting point in the development of the conditioning models for each of the participating jurisdictions.

Because of the large number of these contrasts and the fact that, within each jurisdiction, some contrasts had zero variance, some involved relatively small numbers of individuals, and some were highly correlated with other contrasts or sets of contrasts, an effort was made to reduce the dimensionality of the predictor variables in each jurisdiction's CGROUP model. The original background variable contrasts were standardized and transformed into a set of linearly independent variables by extracting separate sets of principal components (one set for each of the 47 jurisdictions) from the within-state correlation matrices of the original contrast variables. The principal components, rather than the original variables, were used as the independent variables in the conditioning model. The number of principal components included for each jurisdiction was the number required to account for approximately 90 percent of the variance in the original contrast variables. Research based on data from the 1990 Trial State Assessment suggests that results obtained using such a subset of the components will differ only slightly from those obtained using the full set (Mazzeo, Johnson, Bowker, & Fong, 1992).

Table 9-15 contains a listing of the number of principal components included in and the proportion of scale score variance accounted for by the conditioning model for each of the 47 participating jurisdictions. It is important to note that the proportion of variance accounted for by the conditioning model differs across scales within a jurisdiction and across jurisdictions within a scale as summarized in the table. Such variability is not unexpected for at least two reasons. First, there is no reason to expect the strength of the relationship between proficiency and demographics to be identical across all jurisdictions. In fact, one of the reasons for fitting separate conditioning models is that the strength and nature of this relationship may differ across jurisdictions. Second, the homogeneity of the demographic profile also differs across jurisdictions. As with any correlational analysis, the restriction of the range in the predictor variables will attenuate the relationship. However, the proportions are high, falling below .6 for any scale for only one jurisdiction.

Tables 9-16 provides a matrix of estimated within-state correlations among the three fields of science scales averaged over the 47 jurisdictions. In parentheses underneath each average correlation are listed the lowest and highest estimated correlation among the 47 jurisdictions. The listed values, taken directly from the revised CGROUP program, are estimates of the within-state correlations *conditional on the set of principal components included in the conditioning model*.

The number and nature of the scales that were produced were consistent with the recommendations for reporting that were given by the National Assessment Planning Project (see Chapter 2). Reporting results on multiple scales is typically most informative when each of the scales provides unique information about the profile of knowledge and skills possessed by the students being assessed. In such cases, one would hope to see relatively low correlations among the scales. However, with a couple of exceptions, the correlations among the 1996 science scales are high across all jurisdictions, always exceeding .8 and usually exceeding .9. This is particularly noteworthy when one considers that these are correlations *conditional* on a rather large set of background variables. The *marginal* correlations between scales would be higher, particularly for those correlations in the .8 range. In particular, the correlation between two of the scales — *earth science* and *physical science* — are extremely high (rarely falling below .88). The estimated correlations between *life science* and the other two scales are almost as high.

Table 9-15
Proportion of Scale Score Variance Accounted for by Conditioning Models

Jurisdiction	Number of Principal Components	Earth Science	Physical Science	Life Science
Alabama	255	0.74	0.74	0.74
Alaska	202	0.69	0.73	0.69
Arizona	253	0.69	0.72	0.70
Arkansas	240	0.72	0.74	0.72
California	286	0.70	0.73	0.70
Colorado	288	0.66	0.68	0.65
Connecticut	288	0.74	0.74	0.75
Delaware	225	0.71	0.74	0.74
DoDEA/DDESS	153	0.81	0.72	0.76
DoDEA/DoDDS	226	0.61	0.62	0.62
District Of Columbia	198	0.71	0.77	0.74
Florida	301	0.70	0.68	0.74
Georgia	297	0.71	0.71	0.74
Guam	189	0.69	0.73	0.68
Hawaii	201	0.62	0.67	0.66
Indiana	254	0.68	0.70	0.73
Iowa	242	0.64	0.66	0.68
Kentucky	262	0.66	0.64	0.67
Louisiana	280	0.72	0.70	0.72
Maine	234	0.65	0.61	0.67
Maryland	269	0.76	0.76	0.75
Massachusetts	282	0.70	0.70	0.71
Michigan	273	0.72	0.68	0.73
Minnesota	260	0.65	0.68	0.69
Mississippi	273	0.73	0.73	0.72
Missouri	270	0.72	0.72	0.71
Montana	212	0.65	0.68	0.67
Nebraska	254	0.65	0.66	0.69
Nevada	188	0.74	0.76	0.78
New Jersey	223	0.68	0.66	0.69
New Mexico	241	0.73	0.73	0.70
New York	267	0.69	0.69	0.69
North Carolina	262	0.76	0.74	0.77
North Dakota	280	0.67	0.65	0.72
Oregon	219	0.62	0.64	0.69
Pennsylvania	256	0.65	0.68	0.69
Rhode Island	235	0.69	0.71	0.69
South Carolina	268	0.71	0.72	0.74
Tennessee	250	0.73	0.72	0.70
Texas	277	0.77	0.74	0.72
Utah	266	0.64	0.66	0.64

(continued)

Table 9-15 (continued)
Proportion of Scale Score Variance Accounted for by Conditioning Models

Jurisdiction	Number of Principal Components	Earth Science	Physical Science	Life Science
Vermont	217	0.67	0.65	0.71
Virginia	298	0.74	0.75	0.73
Washington	266	0.64	0.64	0.67
West Virginia	263	0.58	0.58	0.65
Wisconsin	268	0.67	0.71	0.72
Wyoming	224	0.66	0.60	0.67

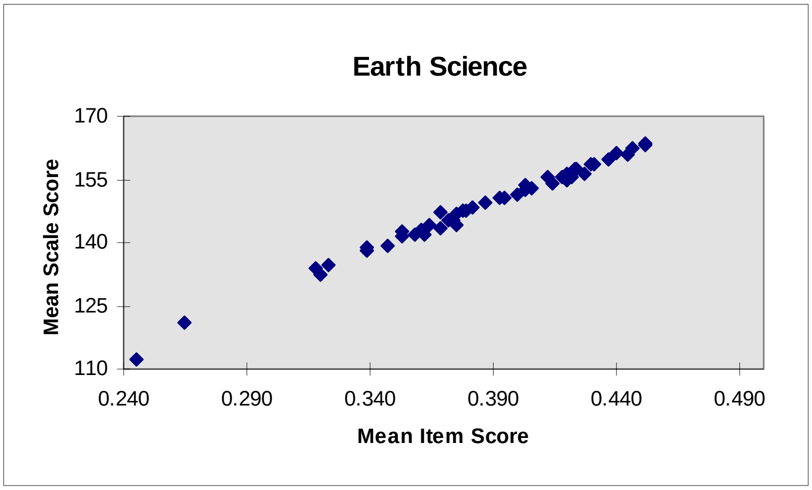
Table 9-16
Average Correlations and Ranges of Scale Correlations Among the Science Scales for 47 Jurisdictions

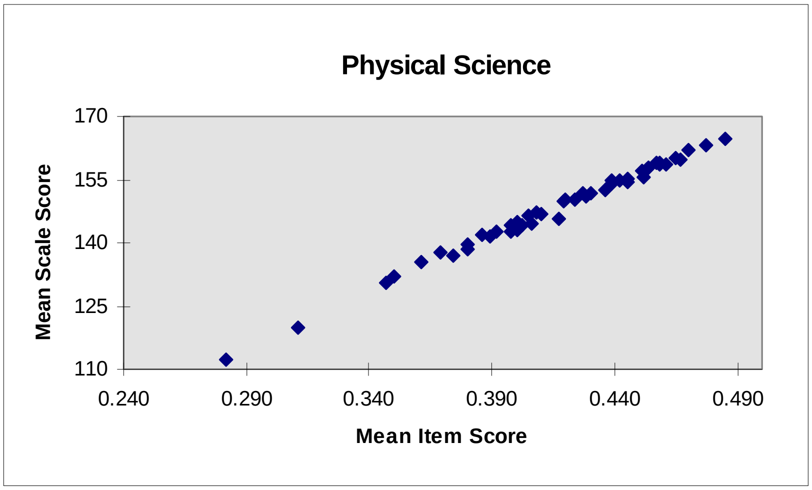
	Earth Science	Physical Science
Earth Science	1.0 (1.0)	0.94 (0.86 - 0.98)
Life Science	0.92 (0.81 - 0.96)	0.93 (0.85 - 0.96)

As discussed in Chapter 8, NAEP scales are viewed as summaries of consistencies and regularities that are present in item-level data. Such summaries should agree with other reasonable summaries of the item-level data. In order to evaluate the reasonableness of the scale score results, a variety of analyses were conducted to compare using scale scores and the average item scores for each fields of science scale. High agreement was found in all of these analyses. One set of such analyses is presented in Figure 9-7. The figure contains scatterplots of the state item score mean versus the state scale score means, for each of the three fields of science scales. As is evident from the plots, there is an extremely strong and almost linear relationship between the estimates of state-level performance in the scale-score and item-score metrics for all three fields of science scales.

Figure 9-7

Plot of Mean Item Score Versus Mean Scale Score for Each Jurisdiction





9.6 LINKING STATE AND NATIONAL SCALES

A major purpose of the State Assessment program was to allow each participating jurisdiction to compare its 1996 results with the nation as a whole and with the region of the country in which that jurisdiction is located.

Although the students in the 1996 State Science Assessment were administered the same test booklets as the eighth graders in the national assessment, separate state and national scalings were carried out (for reasons explained in Mazzeo, 1991, and Yamamoto & Mazzeo, 1992). Again, to ensure a similar scale unit system for the state and national metrics, the scales had to be linked.

For meaningful comparisons to be made between each of the State Assessment jurisdictions and the relevant national samples, results from these two assessments had to be expressed in terms of a similar system of scale units. The purpose of this section is to describe the procedures used to align the 1996 State Assessment scales with their 1996 national counterparts. The procedures that were used represent an extension of the common population equating procedures employed to link the previous national and state scales (Mazzeo, 1991; Yamamoto & Mazzeo, 1992).

Using the house sampling weights provided by Westat (see Section 9.8), the combined sample of students from all participating jurisdictions was used to estimate the distribution of scale scores for the population of students enrolled in public schools that participated in the State Assessment.[5] The total sample size was 97,725. Data from a subsample of the national assessment, consisting of grade-eligible public-school students from any of the 47 jurisdictions that participated in the 1996 State Assessment, was used to obtain estimates of the distribution of scale scores for the same target population. This subsample of national data is referred to as the National Linking sample (NL).[6] Again, appropriate weights provided by Westat were used. Thus, for each of the three scales, two sets of scale score distributions were obtained and used in the linking process. One set, based on the sample of combined data from the State Assessment (referred to as the State Aggregate, or SA, sample) and using item parameter estimates and conditioning results from that assessment, was in the metric of the 1996 State Assessment. The other, based on the sample from the 1996 national assessment (NL) and obtained using item parameters and conditioning results from the national assessment, was in the reporting metric of the 1996 national assessment. The three State Assessment and national scales were made comparable by constraining the mean and standard deviation of the two sets of estimates to be equal.

More specifically, the following steps were followed to linearly link the scales of the two assessments:

1) For each scale, estimates of the scale score distribution for the SA sample was obtained using the full set of plausible values generated by the CGROUP program. The weights used were the final sampling weights provided by Westat (see Section 9.8), not the rescaled versions discussed in Section 9.3. For each scale, the arithmetic mean of the five sets of plausible values was taken as the overall estimated mean and the arithmetic average of the standard deviations of the five sets of plausible values was taken as the overall estimated standard deviation.

---

[5]Students from Guam, DDESS, and DoDDS schools were excluded from the State Aggregate sample sample for purposes of linking .

[6]Note that in previous State Assessments, the National Linking sample was called the State Aggregate Comparison, or SAC, sample. Many people thought this was easy to confuse with state data, so the term ‘National Linking’ will be used in this report.

2) For each scale, the estimated scale score distribution of the NL sample was obtained, again using the full set of plausible values generated by the CGROUP program. The weights used were specially provided by Westat to allow for the estimation of scale score distributions for the same target population of students estimated by the jurisdiction data. The means and standard deviations of the distributions (in the 1996 national reporting metric) for each scale were obtained for this sample in the same manner as described in Step 1.

3) For each scale, a set of linear transformation coefficients was obtained to link the state scale to the corresponding national scale. The linking was of the form

$$Y^* = k_1 + k_2 Y$$

where

Y = a scale score level in terms of the system of units of the provisional BILOG/PARSCALE scale of the State Assessment scaling

Y^* = a scale score level in terms of the system of units comparable to those used for reporting the 1996 national science results

k_2 = [Standard-Deviation$_{NL}$]/[Standard-Deviation$_{SA}$]

k_1 = Mean$_{NL}$ - k_2[Mean$_{SA}$]

where the subscripts refer to the NL sample and to the SA sample.

The final conversion parameters for transforming plausible values from the provisional BILOG/PARSCALE scales to the final State Assessment reporting scales are given in Table 9-17. All State Assessment results are reported in terms of the Y^* metric.

Table 9-17
Transformation Constants for the Scales

Fields of Science Scale	k_1	k_2
Earth Science	34.5966	148.1976
Physical Science	34.1964	148.5352
Life Science	33.5936	148.1027

As is evident from the discussion above, a linear method was used to link the scales from the state and national assessments. While these linear methods ensure equality of means and standard deviations for the SA (after transformation) and the NL samples, they do not guarantee the shapes of the estimated scale score distributions for the two samples to be the same. As these two samples are both from a common target population, estimates of the scale score distribution of that target population based on each of the samples should be quite similar in shape in order to justify strong claims of comparability for the state and national scales. Substantial differences in the shapes of the two estimated distributions would result in differing estimates of the percentages of students above achievement levels or of percentile locations depending on whether State or national scales were used — a clearly unacceptable result given claims about the comparability of the scales. In the face of such results, nonlinear linking methods would be required.

Analyses were carried out to verify the degree to which the linear linking process described above produced comparable scales for State and national results. Comparisons were made between two estimated scale score distributions, one based on the SA sample and one based on the NL sample, for each of the three fields of science scale scales. The comparisons were carried out using slightly modified versions of what Wainer (1974) refers to as suspended rootograms. The final reporting scales for the State and national assessments were each divided into 10-point intervals. Two sets of estimates of the percentage of students in each interval were obtained, one based on the State Assessment aggregate sample and one based on the NL sample. Following Tukey (1971), the square roots of these estimated percentages were compared.[7]

The comparisons are shown in Figures 9-8 through 9-10. The heights of each of the unshaded bars correspond to the square root of the percentage of students from the State Assessment aggregate sample in each 10-point interval on the final reporting scale. The shaded bars show the differences in root percents between the SA and NL estimates. Positive differences indicate intervals in which the estimated percentages from the NL sample are lower than those obtained from the SA. Conversely, negative differences indicate intervals in which the estimated percentages from the NL sample are higher. For all three scales, differences in root percents are quite small, suggesting that the shapes of the two estimated distributions are quite similar (i.e., unimodal with slight negative skewness). There is some evidence that the estimates produced using the SA data are slightly heavier in the extreme lower tails (below 50 for physical science, earth science, and the composite scale; below 30 for life science). However, even these differences at the extremes are small in magnitude (.2 in the root percent metric, .09 in the percent metric) and have little impact on estimates of reported statistics such as percentages of students below the achievement levels.

9.7 PRODUCING A SCIENCE COMPOSITE SCALE

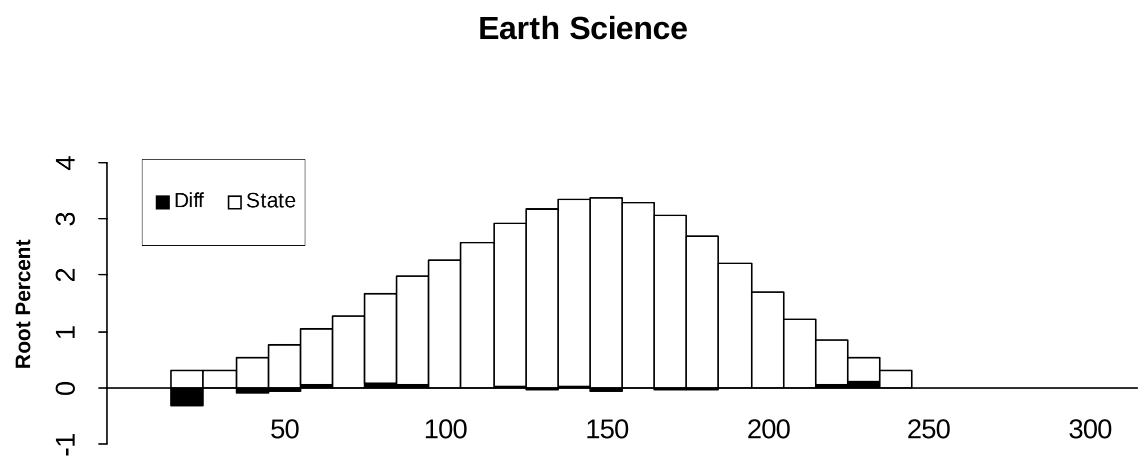
For the national assessment, a composite scale was created for the eighth grade as an overall measure of science scale scores for students at that grade. The composite was a weighted average of plausible values on the three fields of science scale scales (*earth science, physical science, and life science*). The weights for the national fields of science scale scales were proportional to the relative importance assigned to each fields of science scale in each grade in the assessment specifications developed by the Science Objectives Panel. Consequently, the weights for each of the fields of science scales are similar to the actual proportion of items from that fields of science scale.

State Assessment composite scales were developed using weights identical to those used to produce the composites for the 1996 national science assessment. The weights are given in Table 9-18. In developing the State Assessment composite, the weights were applied to the plausible values for each fields of science scale as expressed in terms of the final State Assessment scales (i.e., after transformation from the provisional BILOG/PARSCALE scales.)

[7] The square root transformation allows for more effective comparisons for counts (or equivalently, percentages) when the expected number of counts in each interval is likely to vary greatly over the range of intervals, as is the case for the NAEP scales where the expected counts of individuals in intervals near the extremes of the scale (e.g., below 150 and above 350) are dramatically smaller than the counts obtained near the middle of the scale.

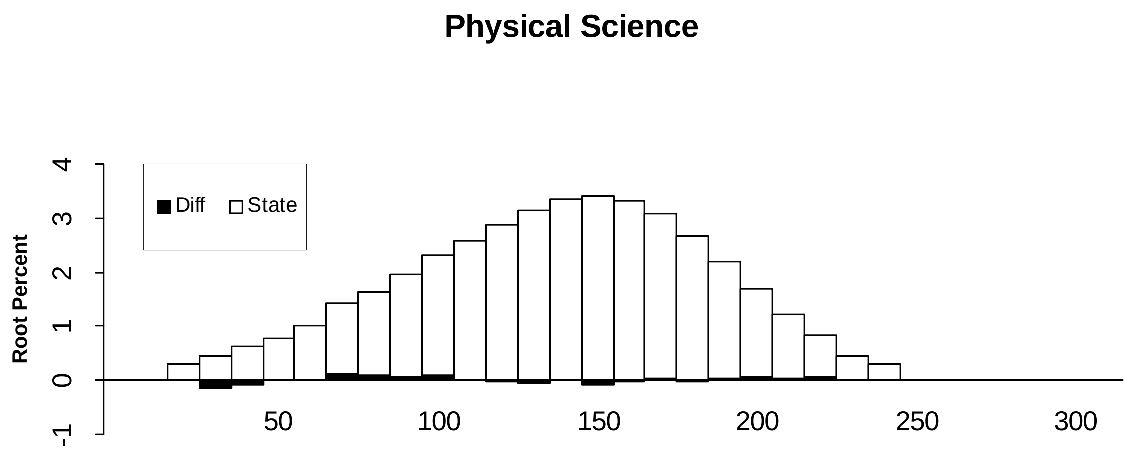
Figure 9-8

Rootogram Comparing Scale Score Distributions for the State Assessment Aggregate Sample and the National Linking Sample for the Earth Science Scale



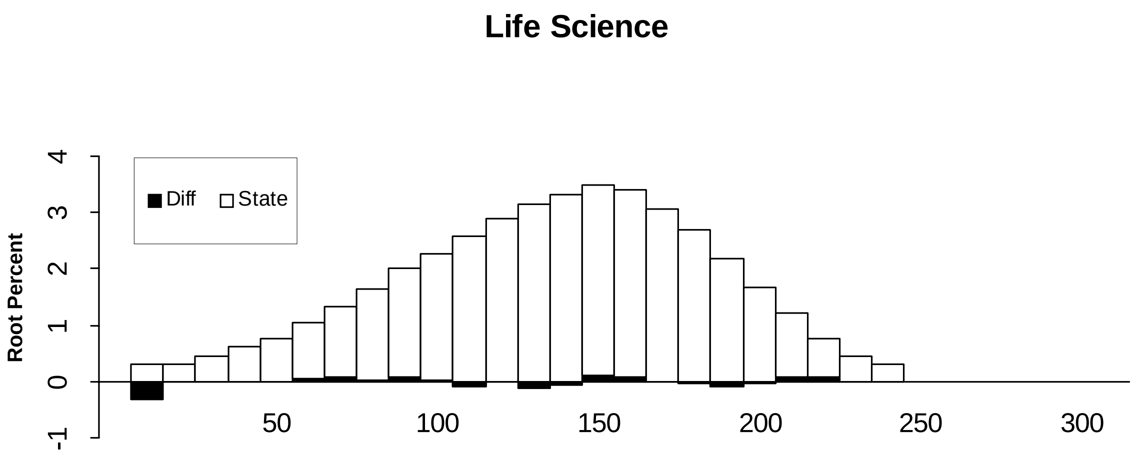
Diff = National - State sample

Figure 9-9
Rootogram Comparing Scale Score Distributions for the State Assessment Aggregate Sample and the National Linking Sample for the Physical Science Scale



Diff = National - State sample

Figure 9-10
Rootogram Comparing Scale Score Distributions for the State Assessment Aggregate Sample and the National Linking Sample for the Life Science Scale



Diff = National - State sample

Table 9-18

Weights Used for Each Scale to Form Composites

Fields of Science Scale	Weight
Earth Science	.30
Physical Science	.30
Life Science	.40

Figure 9-11 provides rootograms comparing the estimated proficiency distributions based on the SA and NL samples for the grade 8 composites. Consistent with the results presented separately by scale, there is some evidence that the estimates produced using the State Assessment data are slightly heavier in the extreme lower tails than the corresponding estimate based on the NL data. However, again these differences in root relative percents are small in magnitude.

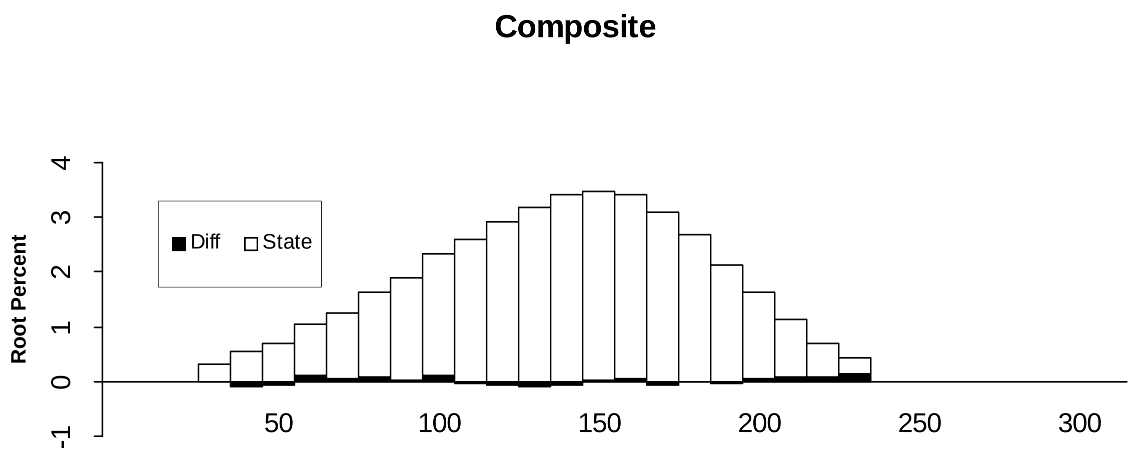
9.8 THE WEIGHT FILES

Westat produced the final student and school weights and the corresponding replicate weights for the 1996 State Assessment. Information for the creation of the weight files was supplied by NCS under the direction of ETS. Because the State Assessment sample was split into two subsamples, one using the 1992 inclusion rules (S1) and one using the 1996 inclusion rules (S2) the weighting process was more complex than in previous assessments. Westat provided two files: a student and school file.

The student weight files contained one record for every student who was not classified as a SD or LEP; the weight files contained two records for every student who was classified as SD or LEP. Each record had a full set of weights, including replicate weights. The first set of weights for the SD and LEP students is to be used when estimating results for either S1 or S2 alone. The second set of weights provided for those students is to be used when estimating results for students from both S1 and S2 together. (See Chapters 3 and 7 for more information about the sampling and weighting procedures for the S1 and S2 samples.)

From the student weight files, ETS constructed three sets of student weights, called modular weights, reporting weights, and all-inclusive weights. The modular weights were used when examining S1 and S2 separately, or for comparing S1 to S2. The reporting weights, used for most reports, were used when reporting results for the students in science who were not classified as being SD or LEP in both S1 and S2 and the students classified as SD or LEP from S2 only. The reporting sample was formed so that valid comparisons with national assessments could be made. (In contrast, for mathematics, only students classified as SD or LEP from S1 were included in the reporting sample.) The SD/LEP students were divided into two types, those who were assessed and those who could not be assessed (called excluded students). The all-inclusive weights were used for estimating results for both S1 and S2 together.

Figure 9-11
Rootogram Comparing Scale Score Distributions for the State Assessment Aggregate Sample and the National Linking Sample for the Composite Scale



Diff = National - State sample

The reporting weights were formed from the student weight files by taking the records for students not classified as SD or LEP, the first record for students in S2 classified as SD or LEP, and a record containing a missing value code for the students in S1 classified as SD or LEP. In this way, the old inclusion rules used with the students classified as SD or LEP in S1 did not effect 1996 State Assessment of science results. For the modular weights, all students not classified as SD or LEP had their final and replicate weights doubled, while the first record for each SD/LEP student was selected directly from the student weight files. It is important to note that the samples should be separated into the S1 and S2 subsamples when using weights generated in this way. To analyze data from S1 and S2 together, the all-inclusive weights should be used. They were created from the student weight files by taking the records for the students not classified as SD or LEP, and the second records for all students classified as SD or LEP.

For the reporting sample, two other weights were created. These are called “house weights” and “senate weights.” As the respective branches of Congress do, these weights represent jurisdictions in two different ways. The house weights weight the student records within a jurisdiction so that the sum of the weights for each jurisdiction is proportional to the fraction of the national in-grade enrollment in that jurisdiction. The senate weights weight the student records within a jurisdiction so that the sums of the weights for each jurisdiction are approximately equal to each other. In other words, a jurisdiction, like California, with many eighth-grade students and a jurisdiction, like Rhode Island, with many fewer eighth-grade students would have equal weight when all of the State Assessment data are combined. Both of these sets of weights are constructed only for the reporting sample. The reporting sample and either the house or senate weights are used during scaling, conditioning and all major reporting.

The house weight is the student’s reporting weight times a factor, which is the number of public school students sampled over the sum of the reporting weights of the public school students in all the jurisdictions. The senate weight is calculated for each jurisdiction separately. Within each jurisdiction a factor, which is 2,000 divided by the sum of the reporting weights of the jurisdiction’s public school students, is computed. (For the 1996 State Assessment, 2,000 rather than the number of public or nonpublic school students within the jurisdiction that was used in previous State Assessments, was used because of the varying sample sizes for each jurisdiction.) The reporting weights for students in both public and nonpublic schools are multiplied by this factor to create the senate weights. For Guam and DoDEA jurisdictions, all schools were considered public in the calculation of these factors.

Each set of weights (modular, reporting, house, senate, and all-inclusive weights) has replicate weights associated with it. Replicate weights are used to estimate jackknife standard errors for each statistic estimated for the State Assessments.

In addition to student weights, school weights are available for use in school level analyses. These weights are modular weights for use when examining S1 and S2 separately or for comparing S1 to S2. No other school weights are available. School level statistics should be calculated on the basis of S1 or S2 subsamples, as opposed to the reporting sample. If school level statistics are calculated for the reporting sample, biases might occur.

9.9 THE GRADE 4 ASSESSMENT OF DDESS AND DoDDS

Because there was no 1996 fourth-grade State Assessment in science, the assessment in DoDEA schools at this grade level required special data analysis and scaling procedures. The five steps mentioned in Section 9.1 were modified to the following three:

- conventional item analysis;
- estimation of proficiency distributions based on the "plausible values" methodology; and
- creation of science composite plausible values.

All analyses were performed treating the DDESS and DoDDS schools as two separate jurisdictions. Item response theory item statistics from the national grade 4 science analysis were used directly in the analysis and their use precluded having to link the DoDEA scales to the national science scales.

Following standard practice in NAEP analyses, the item analyses were carried out in order to check the data. Item statistics were compared to those from the national fourth-grade assessment results, and no data problems were detected.

Using student item response data, data from the background questionnaires (student, teacher, and school) and national item parameters, conditioning model parameters were estimated using the CGROUP computer program, separately for the DDESS and the DoDDS samples. The use of national item parameters was necessary because there was no fourth-grade state assessment and because the two DoDEA samples are not large enough for an independent IRT estimation of item parameters, such as was done for the grade 8 state sample. Also, because there was no fourth-grade State Assessment, it was necessary to use the national parameters. From the resulting conditioning model estimates, the plausible values for each of the science scales were estimated.

These plausible values were transformed to the final science scales (including the composite scale) using the same transformations used with the national fourth-grade plausible values. For each scale other than the composite, the linear transformation obtained for the national grade 4 science scale was of the form:

$$Y^* = k_1 + k_2 Y$$

where

Y = a scale score level in terms of the system of units of the provisional scale of the national assessment scaling (or a DoDEA scale score level)

Y^* = a scale score level in terms of the system of units comparable to those used for reporting the 1996 national science results

k_2 = 35 / (Original National Standard Deviation)

k_1 = 150.0 - k_2 (Original National Mean)

The constants for the three scales are displayed in Table 9-19.

Table 9-19
Transformation Constants

Fields of Science Scale	Original National Mean	Original National Standard Deviation	k_2	k_1
Earth Science	150.0	35	34.0920	150.6685
Physical Science	150.0	35	34.9092	151.1681
Life Science	150.0	35	35.0857	150.5101

The composite scale plausible values were computed as the arithmetic mean of the plausible values on the three scales. This is due to the specification for the grade 4 science instrument that there be an equal number of items in each scale. The plausible values for all scales were then placed on the database for further analysis. Scale score means for various subgroups were computed from the results.

Chapter 10

CONVENTIONS USED IN REPORTING THE RESULTS OF THE 1996 STATE ASSESSMENT PROGRAM IN SCIENCE[1]

Spencer S. Swinton, David S. Freund, and Clyde M. Reese
Educational Testing Service

10.1 OVERVIEW

Results for the 1996 State Assessment in science were disseminated in several different reports: a *Science State Report* for each jurisdiction, a report entitled the *NAEP 1996 Science Report Card for the Nation and the States*, the *Cross-State Data Compendium for the NAEP 1996 Science Assessment*, and, distributed only in electronic form, a six-section almanac of data for each jurisdiction.

The *Science State Report* is a computer-generated report that provides, for each jurisdiction, science results for its eighth-grade students. Although national and regional results[2] are included for comparison purposes, the major focus of each of these computer-generated reports is on the results for a particular jurisdiction. Data about school and student participation rates are reported for each jurisdiction to provide information about the generalizability of the results. School participation rates are reported both in terms of the initially selected samples of schools and in terms of the finally achieved samples, including replacement schools. Several different student participation rates are reported, including the overall rate, the total percentage of students excluded from the assessment, and the exclusion rates for students who are identified as being of limited English proficiency (LEP) and for students with disabilities (SD).

The *State R*eport text and tables were produced by a computerized report generation system developed by ETS report writers, statisticians, data analysts, graphic designers, and editors. Detailed technical documentation about the NAEP computer-generated reporting system can be found in the technical documentation of *The NAEP Computer-Generated Reporting System for the 1994 Trial State Assessment* (Jerry, 1995). Additional information is provided in Section 10.5.3. The reports contain state-level estimates of scale score means and selected percentiles. These results are presented for the state as a whole and for subgroups defined by six key reporting variables (referred to here as primary reporting variables) — gender, race/ethnicity, level of parents' education, Title I participation, eligibility for free or reduced cost

[1] Spencer S. Swinton played a role in making decisions about hypothesis testing methods and procedures and worked with David S. Freund who implemented many of the methods and procedures in computer programs. Spencer and David worked with Clyde M. Reese and others to make decisions about the rules used in generating state reports. John Mazzeo contributed significantly to previous versions of this chapter.

[2]The national and regional results included in the state reports and in portions of the *Cross-State Data Compendium for the NAEP 1996 Science Assessment* are based on data from the 1996 national science assessment and include eighth-grade students enrolled in public and nonpublic schools. Included as public-school jurisdictions in the national sample are eighth-graders enrolled in Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) and Department of Defense Dependents Schools (DoDDS). Fourth-grade students in DDESS and DoDDS schools were also assessed by special arrangement. The results for these two groups of fourth-grader students are not included in the state report series.

school lunch, and type of location. For jurisdictions that secured a sufficient level of participation (see Appendix B), means and percentile results were also reported for students in nonpublic schools (Catholic schools, other religious schools, and other private schools), and for the total in-school population (public-school students, nonpublic-school students, students from the Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) and Department of Defense Dependents Schools (DoDDS), and students attending Bureau of Indian Affairs (BIA) schools). In addition, for public-school students, scale score means were reported for a variety of other subpopulations defined by responses to items from the student, teacher, and school questionnaires and by school and location demographic variables provided by Westat, Inc.[3]

The second report, the *NAEP 1996 Science Report Card for the Nation and the States*, highlights key assessment results for the nation and summarizes results across the jurisdictions participating in the assessment. This report contains composite scale score results (scale score means, etc.) for the nation, for each of the four regions of the country, and for public-school students within each jurisdiction participating in the State Assessment, both overall and by the primary reporting variables.

The third type of report is entitled *Cross-State Data Compendium for the NAEP 1996 Science Assessment*. Like the *Report Card*, the *Compendium* reports results for the nation and for all of the jurisdictions participating in the State Assessment. The *Compendium* contains most of the tables included in the *Report Card* and *State Report* plus tables that provide composite scale results for a large number of secondary reporting variables (e.g., amount of homework, teacher preparation).

The fourth type of summary report is an electronically-delivered almanac that contains a detailed breakdown of the science scale score data according to the responses to the student, teacher, and school questionnaires for the public-school, nonpublic-school, and combined populations as a whole and for important subgroups of the public-school population, as defined by the primary reporting variables. There are six sections to each almanac:

> *The Distribution Data Section* provides selected percentiles for the public-school, nonpublic-school, and total populations and for the major demographic subgroups of the public-school population for the composite scale and each science scale.
>
> *The Student Questionnaire Section* provides a breakdown of the composite scale score data according to the students' responses to questions in the three student questionnaires (common core, science background, and motivational section) included in the assessment booklets.
>
> *The Teacher Questionnaire Section* provides a breakdown of the composite scale score data according to the teachers' responses to questions in the science teacher questionnaire.

---

[3]Some of these variables were used by Westat, Inc., in developing the sampling frame for the assessment and in drawing the sample of participating schools.

The School Questionnaire Section provides a breakdown of the composite scale score data according to the principals' (or other administrators') responses to questions in the school characteristics and policies questionnaire.

The Scale Section provides a breakdown of the scale score data for the three fields of science (*earth science, physical science, or life science*) according to selected items (such as the amount of science homework done per day) from the questionnaires.

The Science Item Section provides the response data (percent of students choosing each option) for each science item in the assessment.

The production of the state reports, the *Report Card*, the *Cross-State Data Compendium*, and the almanacs required many decisions about a variety of data analysis and statistical issues. For example, given the sample sizes obtained for each jurisdiction, certain categories of the reporting variables contained limited numbers of examinees. A decision was needed as to what constituted a sufficient sample size to permit the reliable reporting of subgroup results, and which, if any, estimates were sufficiently unreliable to need to be identified (or flagged) as a caution to readers. As a second example, the state report contained computer-generated text that described the results for a particular jurisdiction and compared total and subgroup performance within the jurisdiction to that of the nation. A number of inferential rules, based on logical and statistical considerations, had to be developed to ensure that the computer-generated reports were coherent from a substantive standpoint and were based on statistical principles of significance testing. Practical comparison procedures were required to control for Type I errors without paying too large a penalty with respect to the statistical power for detecting real and substantive differences. For most tests, family sizes were not so large that the Bonferroni test exacted too large a penalty in power in exchange for protection from Type I error. For sets of comparisons with very large family sizes, such as the state to all other states, a new multiple comparison criterion, False Discovery Rate or FDR (Benjamini & Hochberg, 1994), was implemented, which controls the *rate* of false rejections (e.g., five false rejections per 100 rejections), rather than controlling the probability of even one such error (Familywise Error Rate, or FWE), as does the Bonferroni procedure.

The purpose of this chapter is to document the major conventions and statistical procedures used in generating the state reports, the *Report Card*, the *Cross-State Data Compendium*, and the almanacs. The principal focus of this chapter is on conventions used in the production of the computer-generated state reports. However, Sections 10.2 to 10.4 contain material applicable to all reports. Additional details about procedures relevant to the *Report Card* and *Cross-State Data Compendium* can be found in the text and technical appendices of those reports. Specific guidelines for the publication and notation of NAEP results can be found in Appendix B.

10.2 MINIMUM SCHOOL AND STUDENT SAMPLE SIZES FOR REPORTING SUBGROUP RESULTS

In all of the reports, estimates of quantities such as composite and scale score means and percentages of students indicating particular levels of background variables (as measured in the student, teacher, and school questionnaires) are reported for the population of students in each jurisdiction and grade, as well as for certain key subgroups of interest. The key subgroups were defined by six primary NAEP reporting variables. Where possible, NAEP reports results for gender, for five racial/ethnic subgroups (White, Black, Hispanic, Asian American/Pacific Islander, and American Indian/Alaskan Native), three types of locations (central cities, urban fringes/large towns, rural/small town areas), four levels of parents' education (did not finish high school, high school graduate, some college, college graduate), Title 1 participation, and eligibility for the free or reduced-cost school lunch component of the National School Lunch Program. However, in some jurisdictions, and for some regions of the country, school and/or student sample sizes were quite small for one or more of the categories of these variables. One would expect results for these subgroups to be imprecisely estimated.

It is common practice in reports generated by statistical agencies to suppress those estimates for which the sampling error is so large that it is determined that no effective use can be made of the estimate, or that the potential for misinterpretation outweighs potential benefits of presenting results. A second, and equally important, consideration is whether the standard error estimate that accompanies a statistic is sufficiently accurate to inform potential readers about the reliability of the statistic. The precision of a sample estimate (be it sample mean or standard error estimate) for a population subgroup from a two-stage sample design (such as was used to select the samples for the State Assessment) is a function of the sample size of the subgroup and of the distribution of that sample across first-stage sampling units (i.e., schools in the case of the State Assessment). Hence, both of these factors were used in establishing minimum sample sizes for reporting.

For results to be reported for any subgroup, a minimum student sample size of 62 was required. This number was obtained by determining the sample size necessary to detect an effect size of 0.5 with a probability of 0.8 or greater.[4] The effect size of 0.5 pertains to the "true" difference in mean scale scores between the subgroup in question and the total eighth-grade public-school population in the jurisdiction, divided by the standard deviation of scale scores in the total population. Furthermore, it was required that the students within a subgroup be adequately distributed across schools to allow for reasonably accurate estimation of standard errors. In consultation with Westat, a decision was reached to publish only those statistics that had standard errors estimates based on five or more degrees of freedom. Slightly different variance estimation procedures were used to obtain estimated standard errors for public- and nonpublic-school statistics (see Chapter 7). These different procedures implied different minimum school sample sizes for public- and nonpublic-school results in order to meet the five degrees of freedom minimum. For public-school statistics, subgroup data were required to come from a minimum of five stratification categories formed for variance estimation. For nonpublic-school statistics, a six-school minimum was required.

It should be noted that the full set of reports includes large numbers of tables that provide estimates of the proportion of the students responding to each category of a secondary reporting variable, as well as the mean scale scores of the students within each category. In several instances, the number of students in a particular category of these background variables was also less than 62 or was clustered within a small number of schools. The same minimum student and

[4]A design effect of 2 was assumed for this purpose, implying a sample design-based variance twice that of simple random sampling. This is consistent with previous NAEP experience (Johnson & Rust, 1992). In carrying out the statistical power calculations when comparing a subgroup to the total group, it was assumed that the total population sample size is large enough to make a negligible contribution to standard errors.

school sample size restrictions were applied in the case of proportions as were used for scale score means.

10.3 ESTIMATES OF STANDARD ERRORS WITH LARGE MEAN SQUARED ERRORS

As noted above, standard errors of mean scale scores, proportions, and percentiles play an important role in interpreting subgroup results and in comparing the performances of two or more subgroups. The jackknife standard errors reported by NAEP are statistics whose quality depends on certain features of the sample from which the estimate is obtained. As discussed in the previous section, in certain cases, typically when the number of students upon which the standard error is based is small or when this group of students come from a small number of participating schools, the mean squared error[5] associated with the estimated standard errors may be quite large. Minimum school and student sample sizes were implemented which suppressed statistics in most instances where such problems existed. However, the possibility remained that some statistics based on sample sizes that exceed the minimum requirements might still be associated with standard errors that were not well estimated. Therefore, in the reports, estimated standard errors for published statistics that are subject to large mean squared errors are followed by the symbol "!".

The magnitude of the mean squared error associated with an estimated standard error for the mean or proportion of a group depends on the coefficient of variation (*CV*) of the estimated size of the population group, denoted as $\hat{N}$ (Cochran, 1977, Section 6.3). The coefficient of variation is estimated by:

$$CV(\hat{N}) = \frac{SE(\hat{N})}{\hat{N}}$$

where $\hat{N}$ is a point estimate of N and $SE(\hat{N})$ is the jackknife standard error (described in Chapter 10 of *The NAEP 1994 Technical Report*, Allen, Kline, & Zelenak, 1996) of $\hat{N}$.

Experience with previous NAEP assessments suggests that when this coefficient exceeds 0.2, the mean squared error of the estimated standard errors of means and proportions based on samples of this size may be quite large. (Further discussion of this issue can be found in Johnson & Rust, 1992.) Therefore, the standard errors of means and proportions for all subgroups for which the coefficient of variation of the population size exceeds 0.2 are followed by "!" in the tables of all reports. These standard errors, and any confidence intervals or significance tests involving these standard errors, should be interpreted with caution. In the *Report Card*, the *Cross-State Data Compendium*, and the almanacs, statistical tests involving one or more quantities that have standard errors so flagged should be interpreted with caution.

---

[5]The mean squared error of the estimated standard error is defined as $\mathscr{E}[\hat{S} - \sigma]^2$, where $\hat{S}$ is the estimated standard error, σ is the "true" standard error, and $\mathscr{E}$ is the expectation, or expected value operator.

10.4 TREATMENT OF MISSING DATA FROM THE STUDENT, TEACHER, AND SCHOOL QUESTIONNAIRES

Responses to the student, teacher, and school questionnaires played a prominent role in all reports. Although the return rate on all three types of questionnaire was high,[6] there were missing data for each type of questionnaire.

The reported estimated percentages of students in the various categories of background variables, and the estimates of the mean scale score of such groups, were based on only those students for whom data on the background variable were available. In the terminology of Little and Rubin (1987), the analyses pertaining to a particular background variable presented in the state reports and the *Cross-State Data Compendium* assume the data are missing completely at random (i.e., the mechanism generating the missing data is independent of both the response to the particular background items and the scale score).

The estimates of proportions and proficiencies based on "missing-completely-at-random" assumptions are subject to potential nonresponse bias if, as may be the case, the assumptions are not correct. The amount of missing data was small (usually, less than 2%) for most of the variables obtained from the student and school questionnaires. For analyses based on these variables, reported results are subject to little, if any, nonresponse bias. However, for particular background items from the student and school questionnaires, the level of nonresponse in certain jurisdictions was somewhat higher. As a result, the potential for nonresponse bias in the results of analyses based on this latter set of background items is also somewhat greater. Background items for which more than 10 percent of the returned questionnaires were missing are identified in questionnaire sections of the almanacs produced for each jurisdiction. Again, results for analyses involving these items should be interpreted with caution.

Missing data for students can also occur because students lack the information provided by their teachers on the teacher's questionnaire. In order to analyze the relationships between teachers' questionnaire responses and their students' achievement, each teacher's questionnaire had to be matched to the students who were taught science by that teacher. If a student could not be matched to a teacher, all teacher questionnaire responses are missing for that student. Table 10-1 provides the percentages of eighth-grade students that were matched to teacher questionnaires in each of the 47 jurisdictions that participated in the State Assessment. The first column presents match rates for public-school students and the second for nonpublic-school students. Lower percentages indicate that there is less certainty about results based on data from the teacher questionnaire for that jurisdiction. Note that these match rates do not reflect the additional missing data due to item-level nonresponse. The amount of additional item-level nonresponse in the returned teacher questionnaires can be found in the almanacs produced for each jurisdiction.

[6]Information about survey participation rates (both school and student), as well as proportions of students excluded by each jurisdiction from the assessment, are given in Appendix B. Adjustments intended to account for school and student nonresponse are described in Chapter 7.

Table 10-1

Weighted Percentage of Eighth-Grade Students Matched to Teacher Questionnaires

Jurisdiction	Public	Nonpublic
Alabama	97	100
Alaska	86	--
Arizona	89	92
Arkansas	97	100
California	93	98
Colorado	97	94
Connecticut	98	100
Delaware	95	90
District of Columbia	93	83
DoDEA/DDESS	95	--
DoDEA/DoDDS	94	--
Florida	96	96
Georgia	98	97
Guam	89	84
Hawaii	95	--
Indiana	97	91
Iowa	95	88
Kentucky	94	90
Louisiana	99	100
Maine	95	100
Maryland	97	100
Massachusetts	97	90
Michigan	97	97
Minnesota	96	92
Mississippi	96	89
Missouri	91	99
Montana	93	100
Nebraska	97	86
Nevada	97	100
New Jersey	98	89
New Mexico	92	68
New York	97	91
North Carolina	96	--
North Dakota	95	84
Oregon	95	75
Pennsylvania	97	94
Rhode Island	99	--
South Carolina	97	--
Tennessee	97	--
Texas	92	100
Utah	94	100
Vermont	98	98
Virginia	97	--
Washington	97	--
West Virginia	95	--
Wisconsin	94	100
Wyoming	95	100

10.5 STATISTICAL RULES USED FOR PRODUCING THE STATE REPORTS

As described earlier, the state reports contain jurisdiction-level estimates of eighth-grade mean proficiencies, proportions of students at or above selected scale points, and percentiles for the jurisdiction as a whole and for the categories of a large number of reporting variables. Similar results are provided for the nation and, where sample sizes permitted, for the region to which each jurisdiction belongs.[7] The state reports were computer-generated. The tables and figures, as well as the text of the report, were automatically tailored for each jurisdiction based on the pattern of results obtained. The purpose of this section is to describe some of the procedures and rules used to produce these individually tailored reports. A detailed presentation is available in the technical documentation of *The NAEP Computer-Generated Reporting System for the 1994 Trial State Assessment* (Jerry, 1995). Some changes were made for the 1996 State Assessment, and the current procedures and rules for State Reports are documented in this chapter.

In the 1996 state reports, the results are presented principally through a sequence of tables containing estimated means, proportions, and percentiles, along with their standard errors. In addition to the tables of results, computer-generated interpretive text is also provided. In some cases, the computer-generated interpretive text is primarily descriptive in nature and reports the total group and subgroup scale score means and proportions of interest. However, some of the interpretive text focuses on interesting and potentially important group differences in science scale scores or on the percentages of students responding in particular ways to the background questions. Additional interpretive text compares state-level results with those of the nation. For example, one question of interest to each jurisdiction is whether, on average, its students performed higher than, lower than, or about the same as students in the nation. Additional interpretive text focuses on patterns of achievement across the fields of science or on the pattern of response to a particular background item in the jurisdiction. For example, do more students report spending 30 minutes or 15 minutes on homework each day?

Rules were developed to produce the computer-generated text for comparisons of results for subgroups and for interpretations of patterns of results. These rules were based on a variety of considerations, including a desire for 1) statistical rigor in the identification of important group differences and patterns of results, and 2) solutions that were within the limitations imposed by the availability of computational resources and the time frame for the production of the report. The following sections describe some of these procedures and rules.

10.5.1 Comparing Means and Proportions for Different Groups of Students

Many of the group comparisons explicitly commented on in the state reports involved mutually exclusive sets of students. One common example of such a comparison is the contrast between the mean composite score in a particular jurisdiction and the mean composite score in the nation. Other examples include comparisons within a jurisdiction of the average scale score for male and female students, White and Hispanic students, students attending schools in central city and urban fringe/large town locations, students who reported watching six or more hours of television each night and students who report watching less than one hour each night.

[7]Because U.S. territories are not classified into NAEP regions, no regional comparisons were provided for Guam. Regional results are also not provided for the DDESS and DoDDS schools.

In the state reports, computer-generated text indicated that means or proportions from two groups were different only when the difference in the point estimates for the groups being compared was statistically significant at an approximate simultaneous α level of 0.05. An approximate procedure was used for determining statistical significance NAEP staff judged to be statistically defensible, as well as being computationally tractable. Although all pairs of levels within a variable were tested and reported in the almanacs, computer-generated text for the state reports was developed for only a subset of these comparisons although the family size was maintained at that of the original tests. For example, text was included to compare the majority ethnic group and each minority group, but text for all possible comparisons of groups was not included, even if some unreported comparisons were significant. The procedure used to make statistical tests is described in the following paragraphs.

Let A_i be the statistic in question (i.e., a mean for group *i*) and let S_{A_i} be the jackknife standard error of the statistic. The computer-generated text in the state report identified the means or proportions for groups *i* and *j* as being different if and only if:

$$\frac{|A_i - A_j|}{\sqrt{S^2_{A_i}(A_i) + S^2_{A_j}(A_j)}} \geq T_{\frac{.05}{2c}}$$

where T_α is the $(1 - \alpha)$ percentile of the *t* distribution with degrees of freedom, *df*, as estimated below, and *c* is the number of related comparisons being tested. See Section 10.5.2 for a more specific description of multiple comparisons. In cases where group comparisons were treated as individual units (for example, comparing overall state results with overall national results, the value of *c* was taken as 1, and the test statistic was approximately equivalent to a standard two-tailed t-test for the difference between group means or proportions from large independent samples with the α level set at 0.05. The degrees of freedom of this t-test is defined by a Satterthwaite (Johnson & Rust, 1992) approximation as follows:

$$df = \frac{\left(\sum_{k=1}^{N} S^2_{A_k}\right)^2}{\sum_{k=1}^{N} \frac{S^4_{A_k}}{df_{A_k}}}$$

where N is the number of subgroups involved, four in this case, and df_{A_k} is as follows:

$$df_{A_k} = \left(3.16 - \frac{2.77}{\sqrt{m}}\right)\left[\frac{\left(\sum_{j=1}^{m}(t_{j_k} - t_k)^2\right)^2}{\sum_{j=1}^{m}(t_{j_k} - t_k)^4}\right]$$

The procedures in this section assume that the data being compared are from independent samples. Because of the sampling design used for the State Assessment, in which both schools and students within schools are randomly sampled, the data from mutually exclusive sets of students within a jurisdiction may not be strictly independent. Therefore, the significance tests employed are, in many cases, only approximate. As described in the Section 10.5.4, another procedure, one that does not assume independence, could have been conducted. However, that procedure is computationally burdensome and resources precluded its application for all the comparisons in the state reports. It was the judgment of NAEP staff that if the data were correlated across groups, in most cases the correlation was likely to be positive. Because, in such instances, significance tests based on assumptions of independent samples are conservative (because the estimated standard error of the difference based on independence assumptions is larger than the more complicated estimate based on correlated groups), the approximate procedure was used for most comparisons.

The procedures described above were used for testing differences of both means *and* nonextreme percents. The approximation for the test for percentages works best when sample sizes are large, and the percentages being tested have magnitude relatively close to 50 percent. Statements about group differences should be interpreted with caution if at least one of the groups being compared is small in size and/or if “extreme” percentages are being compared. Percentages, P, were treated as “extreme” if:

$P < P_{\lim} = \dfrac{200}{N_{EFF} + 2}$, where the effective sample size, $N_{EFF} = \dfrac{P(100 - P)}{(SE_{JK})^2}$, and SE_{JK} is the jackknife standard error of P. This “rule of thumb” cutoff leads to flagging a large proportion of confidence intervals that would otherwise include values < 0 or > 1. Similarly, at the other end of the 0 - 100 scale, a percentage is deemed extreme if $100 - P < P_{lim}$. In either extreme case, the normal approximation to the distribution is a poor approximation, and the value of P was reported, but no standard error was estimated and hence no tests were conducted.

10.5.2 Multiple Comparison Procedures

Frequently, groups (or families) of comparisons were made and were presented as a single set. The appropriate text, usually a set of sentences or a paragraph, was selected for inclusion in the State Report based on the results for the entire set of comparisons. For example, Chapter 2 of the state report contains a section that compared average scale scores for a predetermined group, generally the majority group (in the case of race/ethnicity, for example, White students) to those obtained by other minority groups. The entire set of tests was presented in the almanac for that state. For families of comparisons like these, a Bonferroni procedure (Miller, 1966), controlling the Familywise Error Rate (FWE), was used for determining the value of T_α, where c was the number of contrasts in the set. In this example, c was taken to be the number of minority groups meeting minimum sample size requirements, and each statistical test was consequently carried out at an α level of $0.05/c$.

However, in an attempt to gain greater power, two separate definitions of family size were employed for comparisons in two-way tables. For n levels of a control variable (e.g., ethnicity) and m levels of a comparison variable (e.g., number of hours of homework), the standard Bonferroni family size of $n \times m \times (m\text{-}1)/2$ was used. In addition, when the $m \times (m\text{-}1)/2$ marginal tests yielded a significant difference for a pair of categories of the comparison variable, the n levels of the control variable corresponding to that pair of categories were tested with a family size of n. Significance was reported if either definition of family size met the criterion. Further, 2 x 2 interactions were tested for a $m \times n$ table with t-tests using a family size $n \times (n-1) \times m \times (m-1)/4$. In these cases, a modification due to Hochberg of the standard Bonferroni procedure was employed, in which probabilities associated with outcomes are ordered, and α is divided by an integer which increases from 1 to the family size as successively smaller probabilities are tested. More formally, the Hochberg Stagewise Procedure (Hochberg, 1988) is defined:

Let m be the number of significance tests made (the family size) and let $P_1 < P_2 < ... < P_m$ be the ordered significance levels for the m tests. Let α be the combined significance level. The Hochberg procedure compares P_m with α, P_{m-1} with $\alpha/2, ..., P_j$ with $\alpha/(m-j+1)$, stopping comparisons with the first j such that $P_j < \alpha/(m-j+1)$. All tests associated with $P_1, ..., P_j$ are declared significant, all tests associated with $P_{j+1}, ..., P_m$ are declared nonsignificant.

To compare the jurisdiction in a State Assessment report with the nation and all other participating jurisdictions, as is done in the comparisons of overall scale score maps in the State Assessment reports, as many as 46 different comparisons need to be computed. A potentially more powerful multiple comparison procedure was used to judge significance in this case. The procedure, described by Benjamini and Hochberg (1994) was the procedure chosen. Unlike the Bonferroni procedure that controls the FWE, the procedure controls the expected proportion of falsely rejected hypotheses among all rejections. For example, at the 0.05 level, for every 100 rejections of the null hypothesis, the procedure ensures that no more than five will be expected to be false.

The Benjamini and Hochberg application of the False Discovery Rate (FDR) criterion can be described as follows. Let m be the number of significance tests made and let $P_1 \preceq P_2 \preceq . . . \preceq P_m$ be the ordered significance levels of the m tests, from lowest to highest probability. Let α be the combined significance level desired, usually 0.05. The procedure will compare P_m with α, P_{m-1} with $\alpha(m-1)/m$, . . ., P_j with $\alpha j/m$, stopping the comparisons with the first j such that $P_j \preceq \alpha j/m$. All tests associated with P_1, . . ., P_j are declared significant; all tests associated with P_{j+1}, . . ., P_m are declared nonsignificant.

10.5.3 Comparing Proportions Within a Group

Certain analyses in the state report involved the comparison of proportions. One example was the comparison of the proportion of students who reported that a parent graduated from college to the proportion of students who indicated that their parents did not finish high school to determine which proportion was larger. There are other such proportions of interest in this example, such as the proportion of students with at least one parent graduating from high school but neither parent graduating from college. For these types of analyses, NAEP staff determined that the dependencies in the data could not be ignored.

Unlike the case for analyses of the type described in Section 10.5.1, the correlation between the proportion of students reporting a parent graduated from college and the proportion reporting that their parents did not finish high school is likely to be negative and large. For a particular sample of students, it is likely that the higher the proportion of students reporting "at least one parent graduated from college" is, the lower the proportion of students reporting "neither parent graduated from high school" will be. A negative dependence will result in underestimates of the standard error if the estimation is based on independence assumptions (as is the case for the procedures described in Section 10.5.1). Such underestimation can result in an unacceptably large number of "nonsignificant" differences being identified as significant.

The procedures of Section 10.5.1 were modified for the state report analyses that involved comparisons of proportions within a group. The modification involved using a jackknife method for obtaining the standard error of the difference in dependent proportions. The standard error of the difference in proportions was obtained by first obtaining a separate estimate of the difference in question for each jackknife replicate, using the first plausible value only, then taking the standard deviation of the set of replicate estimates as the estimate. The procedures used for proportions within a group differed from the procedures of Section 10.5.1 only with respect to estimating the standard error of the difference; all other aspects of the procedures were identical.

10.5.4 Statistical Significance and Estimated Effect Sizes

Whenever comparisons were made between groups, an attempt was made to distinguish between group differences that were statistically significant but rather small in a practical sense and differences that were both statistically and practically significant. In order to make such distinctions, a procedure based on estimated effect sizes was used. The estimated effect size for comparing means from two groups was defined as:

$$\textit{estimated effect size} = \frac{|A_i - A_j|}{\sqrt{\dfrac{S^2_{A_i} + S^2_{A_j}}{2}}}$$

where A_i refers to the estimated mean for group i, and S_{A_i} refers to the estimated standard deviation within group i. The within-group estimated standard deviations were taken to be the square root of the average of the variances of the set of five plausible values taken over students for each imputation. They were calculated using overall sampling weights provided by Westat.

The estimated effect size for comparing proportions was defined as:

$|f_i - f_j|$, where $f_i = 2 \arcsin \sqrt{p_i}$, and p_i is the estimated proportion in group i (Cohen, 1977).

For both means and proportions, no qualifying language was used in describing significant group differences when the estimated effect size exceeded 0.1. However, when a significant difference was found but the estimated effect size was less than 0.1, the qualifier *somewhat* was used. For example, if the mean scale score for females was significantly higher than that for males but the estimated effect size of the difference was less than 0.1, females were described as performing *somewhat higher* than males.

10.5.5 Descriptions of the Magnitude of Percentages

Percentages reported in the text of the state reports are sometimes described using quantitative words or phrases. For example, the number of students being taught by teachers with master's degrees in English might be described as "relatively few" or "almost all," depending on the size of the percentage in question. Any convention for choosing descriptive terms for the magnitude of percentages is to some degree arbitrary. The rules used to select the descriptive phrases in the report are given in Table 10-2.

Table 10-2

Rules for Descriptive Terms for the Magnitude of Percentages Used in State Reports

Percentage	Description of Text in Report
$p = 0$	None
$0 < p \leq 8$	A small percentage
$8 < p \leq 12$	Relatively few
$12 < p \leq 18$	Less than one fifth
$18 < p \leq 22$	About one fifth
$22 < p \leq 27$	About one quarter
$27 < p \leq 30$	Less than a third
$30 < p \leq 36$	About one third
$36 < p \leq 47$	Less than half
$47 < p \leq 53$	About half
$53 < p \leq 64$	More than half
$64 < p \leq 70$	About two thirds
$70 < p \leq 79$	About three quarters
$79 < p \leq 89$	A large majority
$89 < p < 100$	Almost all
$p = 100$	All

Appendix A

PARTICIPANTS IN THE OBJECTIVES AND ITEM DEVELOPMENT PROCESS

The National Assessment of Educational Progress extends its deep appreciation to all those individuals who participated in the development of the framework, objectives, and items for the 1996 State Assessment program in science.

Project Steering Committee

William O. Baker	Retired, AT&T Bell Laboratories, Murray Hill, New Jersey
Mary Louise Bellamy	Education Director, National Association of Biology Teachers, Reston, Virginia
Frank Betts	Director, Curriculum Technology Center, Association for Supervision and Curriculum Development, Alexandria, Virginia
William B. Campbell	Executive Director, National Industry Council for Science Education, College Park, Maryland
Glenn A. Crosby	Professor, Chemistry Department, Washington State University, Pullman, Washington
Gerald Difford	Executive Director, Colorado Association of Schools, Englewood, Colorado
Janice Earle	Director, Center on Educational Equity, National Association of State Boards of Education, Alexandria, Virginia
John Fowler	Director, Triangle Coalition for Science and Technology, College Park, Maryland
Johnnie Hamilton	Principal, Franklin Intermediate School, Chantilly, Virginia
Elam Hertzler	Retired, Secretary’s Commission on Achieving Necessary Skills, Washington, DC
Ann Kahn	Past President, Parent Teachers Association, Fairfax, Virginia
Douglas Lapp	Director, National Science Resources Center, Smithsonian Institution, Washington, DC
John Layman	Director, Science Teaching Center, University of Maryland, College Park, Maryland

Harold Pratt	Executive Director, Science and Technology Management, Jefferson County Schools, Golden, Colorado
Judith Torney-Purta	Department of Human Development, University of Maryland, College Park, Maryland
Douglas Reynolds	Chief, Bureau of Science Education, State Department of Education, Albany, New York
Bella Rosenberg	Assistant to the President, American Federation of Teachers, Washington, DC
Jane Sisk	Biology Teacher, Calloway County High School, Murray, Kentucky

Project Planning Committee

Andrew Ahlgren	Associate Director, Project 2061, American Association for the Advancement of Science, Washington, DC
Bill Aldridge	Executive Director, National Science Teachers' Association, Washington, DC
J. Myron Atkin	Professor, School of Education, Stanford University, Stanford, California
Joan Boykoff Baron	Assessment Coordinator, Connecticut Common Core of Learning, Bureau of Evaluation and Student Assessment, Connecticut Department of Education, Hartford, Connecticut
Audrey Champagne	Professor, School of Education, State University of New York, Albany, New York
Sally Crissman	Lowers School Head/Science Teacher, Shady Hill School, Cambridge, Massachusetts
Edmund W. Gordon	Professor, Department of Psychology, Yale University, New Haven, Connecticut
Henry Heikkinen	Director, M.A.S.T. Center, University of Northern Colorado, Greeley, Colorado
George Hein	Professor, Lesley College, Cambridge, Massachusetts
Joseph L. Premo	Science Consultant, Minneapolis Public Schools, Minneapolis, Minnesota
Senta A. Raizen	Director, National Center for Improving Science Education, Washington, DC

James Robinson	Retired, Curriculum and Evaluation, Boulder Valley Schools, Boulder, Colorado
Thomas P. Sachse	Manager, Math and Science Unit, California Department of Education, Sacramento, California
Gary E. Skaggs	Test Development Analyst, Office of Research and Evaluation, Fairfax County Public Schools, Falls Church, Virginia

Achievement Levels Panel

Andrew Ahlgren	Associate Director, Project 2061, American Association for the Advancement of Science, Washington, DC
Audrey Champagne	Professor, School of Education, State University of New York, Albany, New York
Richard C. Clark	Consensus Coordinator and Science Coordinator, Minnesota Department of Education, St. Paul, Minnesota
Yvonne Curbeam	Science Department Chair, Dunbar High School, Baltimore, Maryland
Joseph Premo	Science Consultant, New Hope, Minnesota
Thomas Preston	Science Teacher, Frick International Studies Academy, Pittsburgh, Pennsylvania
Senta Raizen	Director, National Center for Improving Science Education, Washington, DC
Dwight Sieggreen	Science Teacher, Cooke Middle School, Northville, Michigan
William Spooner	Chief Consultant, Science Education, North Carolina Department of Public Instruction, Raleigh, North Carolina
Douglas Wagner	Science Teacher, Emmanuel Lutheran School, St. Charles, Missouri
Sylvia Ware	Director, Education Department, American Chemical Society, Washington, DC

Science Project Staff

Ramsay W. Selden	Director, State Education Assessment Center, Council of Chief State School Officers
Richard C. Clark	Consensus Coordinator, Council of Chief State School Officers

Senta A. Raizen	Director, National Center for Improving Science Education
Julia H. Mitchell	Associate Research Scientist, American Institutes for Research
Bonnie L. Verrico	Administrative Assistant, Council of Chief State School Officers

Appendix B

SUMMARY OF PARTICIPATION RATES

Guidelines for Sample Participation and Explanation of the Derivation of Weighted Participation Rates for the 1996 State Assessment Program in Science

Keith F. Rust
Westat, Inc.

and

Eugene G. Johnson and Nada Ballator
Educational Testing Services

Introduction

Since 1989, state representatives, the National Assessment Governing Board (NAGB), several committees of advisors external to the National Assessment of Educational Progress (NAEP), and the National Center for Education Statistics (NCES) have engaged in numerous discussions about the procedures for reporting the NAEP state assessment results. From the outset of these discussions, it was recognized that sample participation rates across the jurisdictions have to be uniformly high to permit fair and valid comparisons. Therefore, NCES established guidelines for school and student participation for the first two Trial State Assessment programs in 1990 and 1992.

The 1994 Trial State Assessment program used an expanded set of participation guidelines. The guidelines were expanded in two ways. First, new guidelines were designed to preempt publication of results from jurisdictions for which participation rates were low enough to suggest the possibility of appreciable nonresponse bias. The new guidelines were congruent both with NAGB policies as well as the resolutions of the Education Information Advisory Committee (EIAC). Second, existing guidelines were extended to cover the presence of separate public- and nonpublic-school samples in the 1994 Trial State Assessment.

For the NAEP 1996 State Assessment program, the participation guidelines implemented in 1994 were again applied. This appendix provides:

- **Participation rate information for the NAEP 1996 State Assessment of science at Grade 8 for both public- and nonpublic-school samples.** This information will also appear in appendices in the *NAEP 1996 Science Report Card* and the *NAEP 1996 Science State Report*.

- **An explanation of the guidelines and notations used in 1996**. In brief, the guidelines cover levels of school and student participation, both overall and for particular population classes, separately for both public- and nonpublic-school samples. Consistent with the NCES standards, weighted data is used to calculate all participation rates for sample surveys, and weighted rates are provided in the reports. The procedures used to derive the weighted school and student participation rates are provided immediately after the discussion of the guidelines and notations.

- **A set of tables that provides the 1996 participation rate information for the NAEP 1996 State Assessment of Science.** Separate information is provided for the public- and nonpublic-school samples. The sample for nonpublic schools includes schools not directed by traditional local or state government agencies, such as those administered by Catholic dioceses, other religious and nonsectarian schools, schools administered by the Bureau of Indian Affairs (BIA) and by the Department of Defense in the United States. Because the aggregate sample across all participating jurisdictions (public or nonpublic) is not necessarily representative of the nation, the weighted participation rates across participating jurisdictions have not been analyzed. However, the counts from the national assessment have been included to provide some context for interpreting the summary of activities in each individual state and territory and for each type of school. Please note that in the NAEP 1996 state assessment, Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) and the Department of Defense Dependents Schools (DoDDS) were included as two separate jurisdictions; for this report and any future reports including 1996 State Assessment data, the two Department of Defense Education Activity (DoDEA) jurisdictions will be reported as jurisdictions having only public schools.

Notations for Use in Reporting School and Student Participation Rates

Unless the overall participation rate is sufficiently high for a jurisdiction, there is a risk that the assessment results for that jurisdiction are subject to appreciable nonresponse bias. Moreover, even if the overall participation rate is high, there may be significant nonresponse bias if the nonparticipation that does occur is heavily concentrated among certain types of schools or students. 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 conditions that will result in the publication of a jurisdiction’s results are presented below. Also presented below are the conditions that will result in a jurisdiction receiving a notation in the 1996 reports. Note that in order for a jurisdiction’s results to be published with no notations, that jurisdiction must satisfy all guidelines.

Guidelines on the Publication of NAEP Results

Guideline 1 - Publication of Public School Results. A jurisdiction will have its public school results published in the *1996 NAEP Science 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 1996 Science 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.

Guideline 2 - Publication of Nonpublic School Results. A jurisdiction will have its nonpublic school results published in the *1996 NAEP Science 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 nonpublic schools is greater than or equal to 70 percent and meets minimum sample size requirements.[1] A jurisdiction eligible to receive a separate *NAEP 1996 Science State Report* under Guideline 1 will have its nonpublic school results included in that report if and only if that jurisdiction's weighted participation rate for the initial sample of nonpublic schools is greater than or equal to 70 percent and meets minimum sample size requirements. If a jurisdiction meets Guideline 2 but fails to meet Guideline 1, a separate State Report will be produced containing only nonpublic school results.

Guideline 3 - Publication of Combined Public and Nonpublic School Results. A jurisdiction will have its combined results published in the *1996 NAEP Science Report Card* (or in other reports that include all state-level results) if and only if both Guidelines 1 and 2 are satisfied. Similarly, a jurisdiction eligible to receive a separate *NAEP 1996 Science State Report* under Guideline 1 will have its combined results included in that report if and only if Guideline 2 is also met.

Discussion. If a jurisdiction's public or nonpublic 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 that 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. Guidelines 1, 2, and 3 take this into consideration. These guidelines are 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.

Guidelines on Notations of NAEP Results

Guideline 4 - 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.

Guideline 5 - Notation for Overall Nonpublic School Participation Rate. A jurisdiction that meets Guideline 2 will receive a notation if its weighted participation rate for the initial sample of nonpublic schools was below 85 percent and the weighted nonpublic school participation rate after substitution was below 90 percent.

[1]Minimum sample size requirements for reporting nonpublic school data consist of two components: (1) a school sample size of six or more participating schools and (2) an assessed student sample size of at least 62.

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 1996 State Assessments, NAEP provided substitutes for nonparticipating public and nonpublic 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% or higher prior to substitution or 90% or higher after substitution), there will be no notation for the relevant overall school participation rate.

Guideline 6 - Notation for Strata-Specific Public School Participation Rates. A jurisdiction that is not already receiving a notation under Guideline 4 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 five 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.

Guideline 7 - Notation for Strata-Specific Nonpublic School Participation Rates. A jurisdiction that is not already receiving a notation under Guideline 5 will receive a notation if the sample of nonpublic 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 five percent of the jurisdiction's total weighted sample of nonpublic schools. The classes of schools from each of which a jurisdiction needed minimum school participation levels were determined by type of nonpublic school (Catholic versus non-Catholic) and location (metropolitan versus nonmetropolitan).

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 minority enrollment, degree of urbanization, and/or median household income, as appropriate for each jurisdiction. For nonpublic schools, nonresponse adjustment cells are determined by type and location of school.

If the weighted response rate, after substitution, for a single adjustment cell falls below 80 percent, and more than five percent (weighted) of the sampled schools are nonparticipants from such a cell, the potential for nonresponse bias is too great. This criterion, that the nonparticipating schools in a class constitute more than five percent of the jurisdiction's population (i.e., the total weighted samples of public or nonpublic schools), is included to insure that the notation is triggered only if the subgroup having a response below 80 percent makes up at least one quarter of the jurisdiction's student population (calculated separately by public and nonpublic schools). This means that a notation is triggered only in cases where a substantial subgroup has experienced an unsatisfactory response. Without this criterion, it is possible that a response rate of just below 80 percent in a small population subgroup could trigger a notation inappropriately. These guidelines are based on the NCES standard for stratum-specific school response rates.

Guideline 8 - 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.

Guideline 9 - Notation for Overall Student Participation Rate in Nonpublic Schools. A jurisdiction that meets Guideline 2 will receive a notation if the weighted student response rate within participating nonpublic schools was below 85 percent.

Discussion. These guidelines follow 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 10 - Notation for Strata-Specific Student Participation Rates in Public Schools. A jurisdiction that is not already receiving a notation under Guideline 8 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 five 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 of limited English proficiency (LEP), and the type of assessment session (monitored or unmonitored), as well as school level of urbanization, minority enrollment, and median household income of the area in which the school is located.

Guideline 11 - Notation for Strata-Specific Student Participation Rates in Nonpublic Schools. A jurisdiction that is not already receiving a notation under Guideline 9 will receive a notation if the sampled students within participating nonpublic 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 five percent of the jurisdiction's weighted assessable nonpublic 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 of limited English proficiency (LEP), and the type of assessment session (monitored or unmonitored), as well as type and location of school.

Discussion. These guidelines address 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 five 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. These guidelines are based on the NCES standard for stratum-specific student response rates.

Derivation of Weighted Participation Rates

Weighted School Participation Rates. The weighted school participation rates within each jurisdiction provide the percentages of eighth-grade students in public (or nonpublic) schools who are represented by the schools participating in the assessment, prior to statistical adjustments for school nonresponse.

Two sets of weighted school participation rates are computed for each jurisdiction, one for public schools and one for nonpublic schools. Each set consists of two weighted participation rates. The first is the weighted participation rate for the initial sample of schools. This rate is based only on those schools that were initially selected for the assessment. The numerator of this rate is the sum of the number of students represented by each initially selected school that participated in the assessment. The denominator is the sum of the number of students represented by each of the initially selected schools found to have eligible students enrolled. This includes both participating and nonparticipating schools.

The second is the weighted participation rate after substitution. The numerator of this rate is the sum of the number of students represented by each of the participating schools, whether originally selected or a substitute. The denominator is the same as that for the weighted participation rate for the initial sample. This means that, for a given jurisdiction and type of school, the weighted participation rate after substitution is always at least as great as the weighted participation rate for the initial sample of schools.

In general, different schools in the sample can represent different numbers of students in the jurisdiction's population. The number of students represented by an initially selected school (the school weight) is the eighth-grade enrollment of the school divided by the probability that the school was included in the sample. For instance, a selected school with a fourth-grade enrollment of 150 and a selection probability of 0.2 represents 750 students from that jurisdiction. The number of students represented by a substitute school is the number of students represented by the replaced nonparticipating school.

Because each selected school represents different numbers of students in the population, the weighted school participation rates may differ somewhat from the simple unweighted rates. (The unweighted rates are calculated from the counts of schools by dividing the number of participating schools by the number of schools in the sample with eligible students enrolled.) The difference between the weighted and the unweighted rates is potentially largest in smaller jurisdictions where all schools with eighth-grade students were included in the sample (that is, where no substitutes are available). In those jurisdictions, each school represents only its own students. Therefore, the nonparticipation of a large school reduces the weighted school participation rate by a greater amount than does the nonparticipation of a small school.

The nonparticipation of larger schools also has greater impact than that of smaller schools on reducing weighted school participation rates in larger jurisdictions where fewer than all of the schools were included in the sample. However, since the number of students represented by each school is more nearly constant in larger states, the difference between the impact of nonparticipation by either large or small schools is less marked than in jurisdictions where all schools were selected.

In general, the greater the population in the jurisdiction, the smaller the difference between the weighted and unweighted school participation rates. However, even in the less populous jurisdictions, the differences tend to be small.

Weighted Student Participation Rate. The weighted student participation rate provides the percentage of the eligible student population from participating schools within the jurisdiction that are represented by the students who participated in the assessment (in either an initial session or a make-up session). Separate weighted student participation rates were calculated for public- and nonpublic-school students. The eligible student population from participating schools (public or nonpublic) within a jurisdiction consists of all students who were in the eighth grade, who attended a school that, if selected, would have participated and who, if selected, would not have been excluded from the assessment. The numerator of this rate is the sum, across all assessed students, of the number of students represented by each assessed student (prior to adjustment for student nonparticipation). The denominator is the sum of the number of students represented by each selected student who was invited and eligible to participate (i.e., not excluded), including students who did not participate. Thus, the denominator is an estimate of the total number of assessable students in the group of schools within the jurisdiction that would have participated if selected.

The number of students represented by a single selected student (the student weight) is 1.0 divided by the overall probability that the student was selected for assessment. In general, the number of students from a jurisdiction's population represented by a sampled student is approximately constant across students. Consequently, there is little difference between the weighted student participation rate and the unweighted student participation rate.

Weighted Overall School and Student Participation Rate. An overall indicator of the effect of nonparticipation by both students and schools is given by the overall participation rate. Separate overall rates were calculated for public- and nonpublic-school samples. For each school type (public or nonpublic), these weights were calculated as the product of the weighted school participation rate (after substitution), and the weighted student participation rate. For jurisdictions having a high overall participation rate, the potential is low for bias to be introduced through either school nonparticipation or student nonparticipation. This rate provides a summary measure that indicates the proportion of the jurisdiction's eighth-grade public or nonpublic school student population that is directly represented by the final student sample. When the overall rate is high, the adjustments for nonresponse that are used in deriving the final survey weights are likely to be effective in maintaining nonresponse bias at a negligible level. Conversely, when the overall rate is relatively low there is a greater chance that a non-negligible bias remains even after making such adjustments.

The overall rate is not used in establishing the guidelines/notations for school and student participation, since guidelines already exist covering school and student participation separately.

Derivation of Weighted Percentages for Excluded Students

Weighted Percentage of Excluded Students. The weighted percentage of excluded students estimates the percentage of the eighth-grade population in the jurisdiction's schools that is represented by the students who were excluded from the assessment, after accounting for school nonparticipation. The numerator is the sum, across all excluded students, of the number of students represented by each excluded student. The denominator is the sum of the number of students represented by each of the students who was sampled (and had not withdrawn from the school at the time of the assessment).

Weighted Percentage of Students with Disabilities (SD). The weighted percentage of SD students estimates the percentage of the eighth-grade population in the jurisdiction's schools represented by the students who were classified as SD, after accounting for school nonparticipation. The numerator is the sum, across all students classified as SD, of the number of students represented by each SD student. The denominator is the sum of the number of students represented by each of the students who was sampled (and had not withdrawn from the school at the time of the assessment).

Weighted Percentage of Excluded SD Students. The weighted percentage of excluded SD students estimates the percentage of students in the jurisdiction who are represented by those SD students excluded from the assessment, after accounting for school nonparticipation. The numerator is the sum, across all students classified as SD and excluded from the assessment, of the number of students represented by each excluded SD student. The denominator is the sum of the number of students represented by each of the students who was sampled (and had not withdrawn from the school at the time of the assessment).

Weighted Percentage of Limited English Proficiency (LEP) Students. The weighted percentage of LEP students estimates the percentage of the eighth-grade population in the jurisdiction's schools represented by the students who were classified as LEP, after accounting for school nonparticipation. The numerator is the sum, across all students classified as LEP, of the number of students represented by each LEP student. The denominator is the sum of the number of students represented by each of the students who was sampled (and had not withdrawn from the school at the time of the assessment).

Weighted Percentage of Excluded LEP Students. The weighted percentage of LEP students who were excluded estimates the percentage of students in the jurisdiction represented by those LEP students excluded from the assessment, after accounting for school nonparticipation. The numerator is the sum, across all students classified as LEP and excluded from the assessment, of the number of students represented by each excluded LEP student. The denominator is the sum of the number of students represented by each student who was sampled (and had not withdrawn from the school at the time of the assessment).

Note: All percentages are based on student weights that have been adjusted for school-level nonresponse. All weighted percentages were calculated separately for public- and nonpublic-school samples.

Table B-1

School Participation Rates, Grade 8, 1996 Science Assessment, Public Schools

Public Schools	Weighted Percentages School Participation Before Substitution	Weighted Percentages School Participation After Substitution	Number of Schools in Original Sample	Number of Schools not Eligible	Number of Schools in Original Sample that Participated	Number of Substituted Schools Provided	Number of Substituted Schools that Participated	Total Number of Schools that Participated
NATION	80	80	155	0	128	3	0	128
JURISDICTIONS								
Alabama	84	90	109	2	90	15	6	96
Alaska[8]	93	93	81	5	55	17	0	55
Arizona	87	87	108	1	94	9	0	94
Arkansas[4]	70	71	111	1	75	25	1	76
California	83	94	108	0	89	19	12	101
Colorado	100	100	110	2	108	0	0	108
Connecticut	100	100	103	1	102	0	0	102
Delaware	100	100	31	1	30	0	0	30
District of Columbia	100	100	36	3	33	0	0	33
DoDEA/DDESS	100	100	11	0	11	0	0	11
DoDEA/DoDDS	100	100	59	1	58	0	0	58
Florida	100	100	110	5	105	0	0	105
Georgia	99	99	107	6	100	1	0	100
Guam	100	100	6	0	6	0	0	6
Hawaii	100	100	52	1	51	0	0	51
Indiana	87	90	107	1	93	12	3	96
Iowa[4]	73	83	115	4	80	27	11	91
Kentucky	87	92	109	1	95	9	5	100
Louisiana	100	100	114	3	111	0	0	111
Maine	91	91	110	4	95	6	0	95
Maryland[6]	86	86	106	2	89	10	0	89
Massachusetts	92	92	107	2	98	7	0	98
Michigan[4]	70	87	107	1	74	31	18	92
Minnesota	86	88	107	0	93	8	2	95
Mississippi	89	95	109	3	96	9	7	103
Missouri	93	96	116	6	102	7	3	105
Montana[4]	70	76	113	5	68	28	11	79
Nebraska	99	100	132	12	119	1	1	120
Nevada[1]	37	38	59	2	27	3	1	28
New Hampshire[1]	66	68	88	0	61	10	3	64
New Jersey[1]	63	64	108	2	66	35	1	67
New Mexico	100	100	90	0	90	0	0	90
New York[4]	70	78	106	0	74	30	8	82
North Carolina	100	100	108	1	107	0	0	107
North Dakota	80	93	125	7	94	19	14	108
Oregon	86	92	111	3	94	12	6	100
Rhode Island	90	90	51	0	43	4	0	43
South Carolina[6]	86	87	107	2	90	10	1	91
Tennessee	92	92	112	4	99	6	0	99
Texas	91	96	109	3	97	9	5	102
Utah	100	100	96	2	94	0	0	94
Vermont[4]	74	75	104	5	77	1	1	78
Virginia	100	100	106	0	106	0	0	106
Washington	94	95	109	0	104	5	1	105
West Virginia	100	100	107	2	105	0	0	105
Wisconsin[4]	78	78	114	0	90	23	0	90
Wyoming	100	100	74	7	67	0	0	67

See preceding text for explanations of the notations and guidelines about sample representativeness and for the derivation of weighted participation. For Delaware, the District of Columbia, DDESS, DoDDS, Guam, Hawaii, and Rhode Island, the State Assessment was based on all eligible public schools (i.e., there was no sampling of public schools).

[1] The state's public school weighted participation rate for the initial sample was less than 70%.

[4] The state's public school weighted participation rate for the initial sample of schools was below 85% and the weighted school participation rate after substitution was below 90%.

[6] The nonparticipating public schools included a class of schools with similar characteristics, which together accounted for more than 5% of the state's total eighth-grade weighted sample of public schools.

[8] The weighted student response rate within participating public schools was below 85%.

In Indiana, Massachusetts, Nebraska, New Hampshire, Oregon, Vermont, and Washington, the materials from one school that conducted an assessment were lost in shipping. The school is included in the counts of participating schools, both before and after substitution. However, in the weighted results, the school is treated in the same manner as a nonparticipating school because no student responses were available for analysis and reporting.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-2

School Participation Rates, Grade 8, 1996 Science Assessment, Nonpublic Schools

Nonpublic Schools	Weighted Percentages School Participation Before Substitution	Weighted Percentages School Participation After Substitution	Number of Schools in Original Sample	Number of Schools not Eligible	Number of Schools in Original Sample that Participated	Number of Substituted Schools Provided	Number of Substituted Schools that Participated	Total Number of Schools that Participated
NATION	77	77	106	6	81	1	0	81
JURISDICTIONS								
Alabama[2]	60	60	19	2	10	7	0	10
Arkansas[5]	74	74	11	2	6	2	0	6
California[5]	80	80	25	6	14	5	0	14
Connecticut[2]	63	65	36	6	19	10	1	20
Delaware[2]	42	44	50	14	12	6	1	13
District of Columbia[2]	52	52	42	8	19	0	0	19
Georgia	88	88	16	5	9	2	0	9
Guam[5]	79	79	10	0	8	0	0	8
Iowa	94	94	22	6	14	2	0	14
Kentucky[5]	82	82	21	4	13	4	0	13
Louisiana[5]	75	75	35	5	21	7	0	21
Maryland[2]	61	64	34	3	18	12	1	19
Massachusetts[5]	75	77	33	5	20	8	1	21
Michigan[5]	80	87	28	3	19	6	2	21
Minnesota[5]	84	84	25	3	19	3	0	19
Missouri	94	100	33	9	23	1	1	24
Montana	93	97	20	5	12	3	1	13
Nebraska[5]	78	84	31	6	17	7	3	20
Nevada	90	90	10	1	8	1	0	8
New Hampshire[5]	83	83	20	4	12	3	0	12
New Jersey[2]	62	64	42	10	20	11	0	20
New Mexico	95	95	21	6	13	2	0	13
New York[5]	84	87	39	6	27	5	1	28
North Dakota[5]	70	78	20	7	9	2	1	10
Oregon[2]	26	26	17	4	4	8	0	4
Rhode Island[2]	68	68	38	5	22	5	0	22
South Carolina[2]	69	69	16	4	8	3	0	8
Texas[5]	79	79	11	1	7	3	0	7
Utah[2]	64	64	10	1	4	3	0	4
Vermont[5]	72	80	23	10	9	2	1	10
Washington	86	86	20	5	11	3	0	11
Wisconsin[2]	65	69	50	8	25	15	2	27
Wyoming[2]	92	92	11	4	6	1	0	6

See preceding text for explanations of the notations and guidelines about sample representativeness and for the derivation of weighted participation.

For the District of Columbia and Guam, the State Assessment was based on all eligible nonpublic schools (i.e., there was no sampling of nonpublic schools).

[2] The state's nonpublic school weighted participation rate for the initial sample was less than 70%.

[5] The state's nonpublic school weighted participation rate for the initial sample of schools was below 85% and the weighted school participation rate after substitution was below 90%.

In the District of Columbia, New Jersey, and Oregon, the materials from one school that conducted an assessment were lost in shipping. The school is included in the counts of participating schools, both before and after substitution. However, in the weighted results, the school is treated in the same manner as a nonparticipating school because no student responses were available for analysis and reporting.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-3

Student Participation Rates, Grade 8, 1996 Science Assessment, Public Schools

Public Schools	Weighted Percentages Student Participation After Make-Ups	Number of Students in Original Sample	Number of Students in Supplemental Sample	Number of Students Withdrawn	Number of Students Excluded	Number of Students To Be Assessed	Number of Students Assessed Initial Sessions	Number of Students Assessed Make-Ups	Total Number of Students Assessed
NATION	93	7,468		153	424	6,891	6,131	245	6,376
JURISDICTIONS									
Alabama	93	2,550	43	138	99	2,356	2,154	32	2,186
Alaska	82	1,931	33	98	45	1,821	1,510	7	1,517
Arizona	90	2,560	94	164	88	2,402	2,112	39	2,151
Arkansas	92	2,120	56	91	75	2,010	1,844	14	1,858
California	92	2,664	63	103	131	2,493	2,251	41	2,292
Colorado	91	2,955	87	192	105	2,745	2,482	32	2,514
Connecticut	93	2,887	32	82	141	2,696	2,426	63	2,489
Delaware	89	2,189	79	109	32	2,127	1,877	26	1,903
District of Columbia	85	2,186	56	139	98	2,005	1,610	90	1,700
DoDEA/DDESS	95	686	35	63	23	635	600	2	602
DoDEA/DoDDS	93	2,562	115	260	41	2,376	2,199	24	2,223
Florida	90	2,812	114	160	153	2,613	2,311	42	2,353
Georgia	92	2,833	74	134	81	2,692	2,442	28	2,470
Guam	90	1,077	45	42	47	1,033	912	18	930
Hawaii	90	2,565	56	166	61	2,394	2,110	43	2,153
Indiana	92	2,665	48	110	87	2,516	2,277	36	2,313
Iowa	94	2,417	38	63	72	2,320	2,152	20	2,172
Kentucky	94	2,701	67	110	60	2,598	2,377	82	2,459
Louisiana	90	3,046	74	132	94	2,894	2,575	40	2,615
Maine	92	2,559	22	33	105	2,443	2,246	8	2,254
Maryland	89	2,482	61	117	76	2,350	2,073	19	2,092
Massachusetts	91	2,640	36	69	118	2,489	2,269	18	2,287
Michigan	90	2,547	44	88	74	2,429	2,137	49	2,186
Minnesota	92	2,699	44	89	57	2,597	2,347	36	2,383
Mississippi	92	2,860	54	133	88	2,693	2,461	8	2,469
Missouri	92	2,746	60	110	89	2,607	2,341	48	2,389
Montana	92	2,305	34	86	47	2,206	2,017	12	2,029
Nebraska	92	3,070	66	91	65	2,980	2,682	42	2,724
Nevada	92	1,112	44	74	34	1,048	959	5	964
New Hampshire	90	1,975	29	48	58	1,898	1,702	8	1,710
New Jersey	93	1,827	26	59	109	1,685	1,565	8	1,573
New Mexico	90	2,870	86	185	143	2,628	2,357	20	2,377
New York	90	2,204	18	41	113	2,068	1,851	25	1,876
North Carolina	91	2,981	66	109	77	2,861	2,593	23	2,616
North Dakota	94	2,692	19	52	18	2,641	2,485	4	2,489
Oregon	89	2,718	84	160	87	2,555	2,235	40	2,275
Rhode Island	89	2,482	37	82	93	2,344	2,083	4	2,087
South Carolina	90	2,523	53	111	79	2,386	2,145	17	2,162
Tennessee	91	2,631	39	118	52	2,500	2,225	62	2,287
Texas	92	2,701	84	162	126	2,497	2,258	42	2,300
Utah	90	3,122	87	142	74	2,993	2,653	62	2,715
Vermont	93	2,149	24	57	66	2,050	1,898	16	1,914
Virginia	90	2,975	63	121	112	2,805	2,538	14	2,552
Washington	90	2,871	54	100	58	2,767	2,482	19	2,501
West Virginia	93	2,984	50	123	110	2,801	2,558	44	2,602
Wisconsin	90	2,506	32	62	106	2,370	2,115	33	2,148
Wyoming	93	2,932	67	139	64	2,796	2,598	21	2,619

See preceding text for explanations of the notations and guidelines about sample representativeness and for the derivation of weighted participation. For the national sample, column 3 is not applicable.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-4

Student Participation Rates, Grade 8, 1996 Science Assessment, Nonpublic Schools

Nonpublic Schools	Weighted Percentages Student Participation After Make-Ups	Number of Students in Original Sample	Number of Students in Supplemental Sample	Number of Students Withdrawn	Number of Students Excluded	Number of Students To Be Assessed	Number of Students Assessed Initial Sessions	Number of Students Assessed Make-Ups	Total Number of Students Assessed
NATION	97	1,461		7	2	1,452	1,382	16	1,398
JURISDICTIONS									
Alabama	95	151	1	2	0	150	144	0	144
Arkansas	99	90	1	1	0	90	89	0	89
California	96	213	2	1	0	214	206	0	206
Connecticut	96	278	0	2	2	274	263	0	263
Delaware	96	325	1	1	0	325	296	17	313
District of Columbia	95	276	0	2	0	274	256	3	259
Georgia	96	244	11	13	0	242	232	0	232
Guam	94	213	1	4	0	210	198	0	198
Iowa	96	255	1	0	0	256	246	0	246
Kentucky	97	269	4	5	0	268	260	0	260
Louisiana	96	448	3	5	1	445	424	0	424
Maryland	94	347	0	4	0	343	318	4	322
Massachusetts	94	360	2	2	3	357	335	0	335
Michigan	97	344	7	4	3	344	318	14	332
Minnesota	94	266	2	4	0	264	247	0	247
Missouri	95	383	5	2	1	385	365	0	365
Montana	93	158	9	3	1	163	154	0	154
Nebraska	96	346	3	2	1	346	332	1	333
Nevada	91	149	2	3	2	146	133	0	133
New Hampshire	95	189	1	2	0	188	179	0	179
New Jersey	96	299	0	1	1	297	287	0	287
New Mexico	95	243	4	6	0	241	229	1	230
New York	97	537	0	4	3	530	513	1	514
North Dakota	93	169	2	2	1	168	160	0	160
Oregon	86	62	2	2	0	62	54	0	54
Rhode Island	96	359	0	4	3	352	340	0	340
South Carolina	95	142	3	1	0	144	138	0	138
Texas	98	132	2	1	0	133	130	0	130
Utah	93	99	2	1	0	100	93	0	93
Vermont	91	131	1	4	1	127	115	0	115
Washington	95	235	1	9	0	227	215	0	215
Wisconsin	96	395	1	2	1	393	380	0	380
Wyoming	94	47	3	0	0	50	47	0	47

See preceding text for explanations of the notations and guidelines about sample representativeness and for the derivation of weighted participation. For the national sample, column 3 is not applicable.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-5

Summary of School and Student Participation, Grade 8
1996 Science Assessment, Public Schools

Public Schools	Weighted Percentages School Participation Before Substitution	Notation Number 1	Weighted Percentages School Participation After Substitution	Notation Number 4	Notation Number 6	Weighted Percentages Student Participation After Make-Ups	Notation Number 8	Weighted Overall Rate
NATION	80		80			93		74
JURISDICTIONS								
Alabama	84		90			93		83
Alaska	93		93			82	*	76
Arizona	87		87			90		78
Arkansas	70		71	*		92		65
California	83		94			92		86
Colorado	100		100			91		91
Connecticut	100		100			93		93
Delaware	100		100			89		89
District of Columbia	100		100			85		85
DoDEA/DDESS	100		100			95		95
DoDEA/DoDDS	100		100			93		93
Florida	100		100			90		90
Georgia	99		99			92		91
Guam	100		100			90		90
Hawaii	100		100			90		90
Indiana	87		90			92		83
Iowa	73		83	*		94		77
Kentucky	87		92			94		87
Louisiana	100		100			90		90
Maine	91		91			92		83
Maryland	86		86		*	89		76
Massachusetts	92		92			91		85
Michigan	70		87	*		90		78
Minnesota	86		88			92		81
Mississippi	89		95			92		87
Missouri	93		96			92		88
Montana	70		76	*		92		70
Nebraska	99		100			92		92
Nevada	37	*	38			92		35
New Hampshire	66	*	68			90		62
New Jersey	63	*	64			93		59
New Mexico	100		100			90		90
New York	70		78	*		90		70
North Carolina	100		100			91		91
North Dakota	80		93			94		88
Oregon	86		92			89		82
Rhode Island	90		90			89		80
South Carolina	86		87		*	90		78
Tennessee	92		92			91		84
Texas	91		96			92		89
Utah	100		100			90		90
Vermont	74		75	*		93		70
Virginia	100		100			90		90
Washington	94		95			90		86
West Virginia	100		100			93		93
Wisconsin	78		78	*		90		71
Wyoming	100		100			93		93

See preceding text for explanations of the notations and guidelines about sample representativeness and for the derivation of weighted participation.

Notation Number 1: The state's public school weighted participation rate for the initial sample was less than 70%.

Notation Number 4: The state's public school weighted participation rate for the initial sample of schools was below 85% and the weighted school participation rate after substitution was below 90%.

Notation Number 6: The nonparticipating public schools included a class of schools with similar characteristics, which together accounted for more than 5% of the state's total eighth-grade weighted sample of public schools.

Notation Number 8: The weighted student response rate within participating public schools was below 85%.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-6

Summary of School and Student Participation, Grade 8
1996 Science Assessment, Nonpublic Schools

Nonpublic Schools	Weighted Percentages School Participation Before Substitution	Notation Number 2	Weighted Percentages School Participation After Substitution	Notation Number 5	Notation Number 7	Weighted Percentages Student Participation After Make-Ups	Notation Number 9	Weighted Overall Rate
NATION	77		77			97		74
JURISDICTIONS								
Alabama	60	*	60			95		57
Arkansas	74		74	*		99		73
California	80		80	*		96		76
Conneticut	63	*	65			96		62
Delaware	42	*	44			96		42
District of Columbia	52	*	52			95		50
Georgia	88		88			96		84
Guam	79		79	*		94		74
Iowa	94		94			96		90
Kentucky	82		82	*		97		79
Louisiana	75		75	*		96		71
Maryland	61	*	64			94		60
Massachusetts	75		77	*		94		73
Michigan	80		87	*		97		84
Minnesota	84		84	*		94		79
Missouri	94		100			95		95
Montana	93		97			93		90
Nebraska	78		84	*		96		81
Nevada	90		90			91		82
New Hampshire	83		83	*		95		79
New Jersey	62	*	64			96		61
New Mexico	95		95			95		91
New York	84		87	*		97		85
North Dakota	70		78	*		93		73
Oregon	26	*	26			86		22
Rhode Island	68	*	68			96		66
South Carolina	69	*	69			95		65
Texas	79		79	*		98		77
Utah	64	*	64			93		60
Vermont	72		80	*		91		73
Washington	86		86			95		81
Wisconsin	65	*	69			96		66
Wyoming	92	*	92			94		87

See preceding text for explanations of the notations and guidelines about sample representativeness and for the derivation of weighted participation.

Notation Number 2: The state's nonpublic school weighted participation rate for the initial sample was less than 70%.

Notation Number 5: The state's nonpublic school weighted participation rate for the initial sample of schools was below 85% and the weighted school participation rate after substitution was below 90%.

Notation Number 7: The nonparticipating nonpublic schools included a class of schools with similar characteristics, which together accounted for more than 5% of the state's total eighth-grade weighted sample of nonpublic schools.

Notation Number 9: The weighted student response rate within participating nonpublic schools was below 85%.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-7

Weighted Percentages of Students Excluded (SD and LEP) from Student Sample
Grade 8, 1996 Science Assessment, Public Schools

Public Schools	Total Percentages Students Identified SD and LEP	Total Percentages Students Excluded	Percentages Students Identified SD	Percentages Students Excluded SD	Percentages Students Identified LEP	Percentages Students Excluded LEP
NATION	14	5	10	4	4	2
JURISDICTIONS						
Alabama	11	7	11	7	0	0
Alaska	16	5	13	4	5	1
Arizona	16	6	9	5	7	2
Arkansas	11	7	10	6	1	1
California	21	9	8	4	14	6
Colorado	12	7	10	5	3	3
Connecticut	15	9	13	8	2	2
Delaware	11	2	10	2	1	1
District of Columbia	12	9	10	7	3	2
DoDEA/DDESS	10	7	8	5	3	3
DoDEA/DoDDS	8	3	6	2	2	1
Florida	18	10	15	8	4	2
Georgia	11	6	10	5	1	1
Guam	11	8	7	5	5	3
Hawaii	11	5	9	4	2	1
Indiana	11	6	11	6	1	0
Iowa	15	6	14	5	1	0
Kentucky	9	4	9	4	0	0
Louisiana	11	6	10	6	0	0
Maine	13	7	13	7	1	0
Maryland	11	5	10	5	2	1
Massachusetts	18	8	15	6	3	2
Michigan	10	5	9	5	1	0
Minnesota	12	4	11	4	2	1
Mississippi	10	6	10	6	0	0
Missouri	13	6	13	6	1	0
Montana	9	3	9	3	0	0
Nebraska	12	4	11	4	1	0
Nevada	9	6	6	4	3	2
New Hampshire	14	5	14	5	0	0
New Jersey	16	11	13	8	4	3
New Mexico	20	9	15	7	7	3
New York	16	9	8	4	8	5
North Carolina	9	5	9	5	1	1
North Dakota	7	2	7	2	0	0
Oregon	12	5	9	4	3	2
Rhode Island	17	7	13	5	5	2
South Carolina	10	6	10	6	0	0
Tennessee	12	4	12	4	1	1
Texas	18	8	11	6	8	3
Utah	9	4	8	4	1	1
Vermont	14	6	13	5	1	1
Virginia	13	7	10	6	4	1
Washington	11	4	8	3	3	1
West Virginia	13	7	13	7	0	0
Wisconsin	12	8	11	7	2	1
Wyoming	11	5	10	4	1	0

SD = Students with Disabilities (the term previously used was IEP).
LEP = Limited English Proficiency.
To be excluded, a student was supposed to be classified as SD or as LEP and judged incapable of participating in the assessment. A student reported as belonging to both SD and LEP classifications is counted once in the overall rate (first column), once in the overall excluded rate (second column), and separately in the remaining columns.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-8

Weighted Percentages of Students Excluded (SD and LEP) from Student Sample Grade 8, 1996 Science Assessment, Nonpublic Schools

Nonpublic Schools	Total Percentage Students Identified SD and LEP	Total Percentage Students Excluded	Percentage Students Identified SD	Percentage Students Excluded SD	Percentage Students Identified LEP	Percentage Students Excluded LEP
NATION	3	0	2	0	0	0
JURISDICTIONS						
Alabama	0	0	0	0	0	0
Arkansas	2	0	2	0	0	0
California	1	0	1	0	0	0
Connecticut	4	1	3	1	1	1
Delaware	1	0	1	0	0	0
District of Columbia	4	0	4	0	0	0
Georgia	0	0	0	0	0	0
Guam	0	0	0	0	0	0
Iowa	1	0	1	0	0	0
Kentucky	0	0	0	0	0	0
Louisiana	5	1	5	1	0	0
Maryland	1	0	1	0	0	0
Massachusetts	5	2	1	0	4	2
Michigan	4	2	3	0	2	2
Minnesota	0	0	0	0	0	0
Missouri	5	0	5	0	0	0
Montana	13	1	1	1	12	0
Nebraska	2	0	1	0	0	0
Nevada	2	2	2	2	0	0
New Hampshire	0	0	0	0	0	0
New Jersey	6	1	6	1	0	0
New Mexico	4	0	4	0	0	0
New York	2	1	2	1	0	0
North Dakota	15	1	6	1	10	1
Oregon	0	0	0	0	0	0
Rhode Island	3	2	3	2	0	0
South Carolina	0	0	0	0	0	0
Texas	4	0	4	0	0	0
Utah	0	0	0	0	0	0
Vermont	1	1	0	0	1	1
Washington	1	0	0	0	1	0
Wisconsin	2	1	1	1	0	0
Wyoming	10	0	10	0	0	0

SD = Students with Disabilities (the term previously used was IEP).
LEP = Limited English Proficiency.
To be excluded, a student was supposed to be classified as SD or as LEP and judged incapable of participating in the assessment. A student reported as belonging to both SD and LEP classifications is counted once in the overall rate (first column), once in the overall excluded rate (second column), and separately in the remaining columns.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-9

Weighted Percentages of Absent, SD, and LEP Students Based on Those Invited to Participate Grade 8, 1996 Science Assessment, Public Schools

Public Schools	Weighted Percentages Student Participation After Make-Ups	Weighted Percentages Absent	Weighted Percentages Assessed SD	Weighted Percentages Absent SD	Weighted Percentages Assessed LEP	Weighted Percentages Absent LEP
NATION	93	7	87	13	93	7
JURISDICTIONS						
Alabama	93	7	89	11	100	0
Alaska	82	18	76	24	90	10
Arizona	90	10	82	18	93	7
Arkansas	92	8	89	11	81	19
California	92	8	86	14	91	9
Colorado	91	9	84	16	90	10
Connecticut	93	7	94	6	100	0
Delaware	89	11	81	19	78	22
District of Columbia	85	15	85	15	100	0
DoDEA/DDESS	95	5	92	8	0	100
DoDEA/DoDDS	93	7	92	8	72	28
Florida	90	10	85	15	85	15
Georgia	92	8	89	11	75	25
Guam	90	10	72	28	67	33
Hawaii	90	10	88	12	83	17
Indiana	92	8	89	11	100	0
Iowa	94	6	93	7	100	0
Kentucky	94	6	90	10	100	0
Louisiana	90	10	80	20	100	0
Maine	92	8	88	12	80	20
Maryland	89	11	81	19	84	16
Massachusetts	91	9	88	12	95	5
Michigan	90	10	75	25	100	0
Minnesota	92	8	84	16	77	23
Mississippi	92	8	89	11	33	67
Missouri	92	8	90	10	100	0
Montana	92	8	82	18	100	0
Nebraska	92	8	90	10	88	12
Nevada	92	8	95	5	100	0
New Hampshire	90	10	85	15	100	0
New Jersey	93	7	96	4	100	0
New Mexico	90	10	83	17	89	11
New York	90	10	84	16	89	11
North Carolina	91	9	95	5	100	0
North Dakota	94	6	91	9	100	0
Oregon	89	11	86	14	75	25
Rhode Island	89	11	79	21	97	3
South Carolina	90	10	85	15	100	0
Tennessee	91	9	89	11	100	0
Texas	92	8	91	9	92	8
Utah	90	10	84	16	84	16
Vermont	93	7	86	14	100	0
Virginia	90	10	77	23	85	15
Washington	90	10	87	13	86	14
West Virginia	93	7	81	19	100	0
Wisconsin	90	10	79	21	80	20
Wyoming	93	7	95	5	92	8

SD = Students with Disabilities (the term previously used was IEP).
LEP = Limited English Proficiency.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-10

Weighted Percentages of Absent, SD, and LEP Students Based on Those Invited to Participate Grade 8, 1996 Science Assessment, Nonpublic Schools

Nonpublic Schools	Weighted Percentages Student Participation After Make-Ups	Weighted Percentages Absent	Weighted Percentages Assessed SD	Weighted Percentages Absent SD	Weighted Percentages Assessed LEP	Weighted Percentages Absent LEP
NATION	97	3	93	7	61	39
JURISDICTIONS						
Alabama	95	5	0	100	0	100
Arkansas	99	1	100	0	0	100
California	96	4	0	100	0	100
Connecticut	96	4	100	0	100	0
Delaware	96	4	100	0	0	100
District of Columbia	95	5	100	0	0	100
Georgia	96	4	0	100	0	100
Guam	94	6	0	100	0	100
Iowa	96	4	100	0	0	100
Kentucky	97	3	0	100	0	100
Louisiana	96	4	100	0	0	100
Maryland	94	6	100	0	0	100
Massachusetts	94	6	100	0	100	0
Michigan	97	3	100	0	0	100
Minnesota	94	6	0	100	0	100
Missouri	95	5	100	0	0	100
Montana	93	7	0	100	80	20
Nebraska	96	4	100	0	0	100
Nevada	91	9	0	100	0	100
New Hampshire	95	5	0	100	0	100
New Jersey	96	4	100	0	0	100
New Mexico	95	5	85	15	0	100
New York	97	3	100	0	0	100
North Carolina	93	7	100	0	82	18
Oregon	86	14	0	100	0	100
Rhode Island	96	4	100	0	0	100
South Carolina	95	5	0	100	0	100
Texas	98	2	100	0	0	100
Utah	93	7	0	100	0	100
Vermont	91	9	0	100	0	100
Washington	95	5	0	100	100	0
Wisconsin	96	4	100	0	100	0
Wyoming	94	6	100	0	0	100

SD = Students with Disabilities (the term previously used was IEP).
LEP = Limited English Proficiency.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment.

Table B-11

Questionnaire Response Rates, Grade 8, 1996 Science Assessment, Public Schools

Public Schools	Weighted Percentage of Students Matched to Science Teacher Questionnaires	Weighted Percentage of Students Matched to School Characteristics/ Policies Questionnaires	Percentage of School Characteristics/ Policies Questionnaires Returned	Percentage of SD/LEP Student[1] Questionnaires Returned
NATION	85	92	89	86
JURISDICTIONS				
Alabama	94	90	90	87
Alaska	76	80	85	59
Arizona	91	90	88	84
Arkansas	96	86	86	93
California	90	94	93	75
Colorado	87	94	94	86
Connecticut	89	79	79	90
Delaware	89	93	93	87
District of Columbia	68	77	76	68
DoDEA/DDESS	86	82	82	95
DoDEA/DoDDS	90	93	97	89
Florida	86	96	96	81
Georgia	91	88	88	91
Guam	81	82	83	49
Hawaii	41	96	94	85
Indiana	89	93	91	83
Iowa	94	97	98	87
Kentucky	84	93	93	90
Louisiana	88	91	91	88
Maine	90	93	94	88
Maryland	91	83	82	91
Massachusetts	89	86	86	92
Michigan	88	85	85	85
Minnesota	91	85	84	91
Mississippi	90	91	91	89
Missouri	88	93	92	85
Montana	90	89	87	89
Nebraska	89	87	91	91
Nevada	88	97	93	96
New Hampshire	99	88	87	95
New Jersey	95	91	90	87
New Mexico	85	81	80	61
New York	83	74	77	78
North Carolina	91	91	91	92
North Dakota	95	93	91	85
Oregon	88	93	92	77
Rhode Island	96	89	88	81
South Carolina	91	92	91	88
Tennessee	91	95	95	91
Texas	92	86	86	84
Utah	87	92	91	88
Vermont	92	90	91	89
Virginia	93	98	97	85
Washington	80	86	88	84
West Virginia	91	91	90	90
Wisconsin	83	91	91	87
Wyoming	86	94	92	86

[1] This percentage is unweighted, and computed over all students, whereas all other numbers in this report are based on the reporting sample only.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment

Table B-12

Questionnaire Response Rates, Grade 8, 1996 Science Assessment, Nonpublic Schools

Nonpublic Schools	Weighted Percentage of Students Matched to Science Teacher Questionnaires	Weighted Percentage of Students Matched to School Characteristics/ Policies Questionnaires	Percentage of School Characteristics/ Policies Questionnaires Returned	Percentage of SD/LEP Student[1] Questionnaires Returned
NATION	91	92	92	95
JURISDICTIONS				
Alabama	79	80	80	0
Arkansas	99	100	100	100
California	83	86	86	100
Connecticut	77	86	85	54
Delaware	79	82	85	100
District of Columbia	88	85	82	100
Georgia	75	64	78	0
Guam	87	75	88	100
Iowa	88	92	93	100
Kentucky	90	87	85	100
Louisiana	85	89	86	100
Maryland	78	96	95	33
Massachusetts	84	92	90	85
Michigan	90	87	86	100
Minnesota	84	98	95	100
Missouri	90	74	79	100
Montana	88	70	77	100
Nebraska	87	93	95	88
Nevada	82	81	88	100
New Hampshire	91	98	92	100
New Jersey	89	97	95	88
New Mexico	94	100	100	100
New York	94	74	71	100
North Dakota	86	100	100	71
Oregon	100	66	67	0
Rhode Island	88	77	82	100
South Carolina	86	89	88	0
Texas	100	100	100	100
Utah	100	100	100	100
Vermont	100	100	100	0
Washington	100	100	100	100
Wisconsin	88	90	89	80
Wyoming	100	49	83	100

[1] This percentage is unweighted, and computed over all students, whereas all other numbers in this report are based on the reporting sample only.

SOURCE: National Center for Education Statistics. National Assessment of Educational Progress (NAEP). 1996 Science Assessment

Appendix C

CONDITIONING VARIABLES AND CONTRAST CODINGS

This appendix contains information about the conditioning variables used in scaling/plausible value estimation for the 1996 State Assessment program. The initial step in construction of conditioning variables involves forming primary student-based vectors of response data from answers to student, teacher, and school questionnaires, demographic and background data such as supplied by Westat, Inc., and other student information known prior to scaling. The initial conditioning vectors concatenate this student background information into a series of identifying “contrasts” comprising:

1. Categorical variables derived by expanding the response options of a questionnaire variable into a binary series of one-degree-of-freedom “dummy” variables or contrasts, (these form the majority of each student conditioning vector);

2. Questionnaire or demographic variables that possess ordinal response options, such as number of hours spent watching television, which are included as linear and/or quadratic multi-degree-of-freedom contrasts;

3. Continuous variables, such as student logit scores based on percent correct values, included as contrasts in their original form or a transformation of their original form, and;

4. Interactions of two or more categorical variables forming a set of orthogonal one-degree-of-freedom dummy variables or contrasts.

This appendix gives the specifications used for constructing the conditioning variables. Table C-1 defines the information provided for each main sample variable. Table C-2 provides a summary of the conditioning variables specifications that are contained in the remainder of this appendix.

As described in Chapter 9, the linear conditioning model employed for the estimation of plausible values did not directly use the conditioning variable specifications listed in this appendix. To eliminate inherent instabilities in estimation encountered when using a large number of correlated variables, a principal component transformation of the correlation matrix obtained from the conditioning variable contrasts derived according to these primary specifications was performed. The principal components scores based on this transformation were used as the predictor variables in estimating the linear conditioning model.

Table C-1

Description of Specifications Provided for Each Conditioning Variable

Title	Description
CONDITIONING ID	A unique eight-character ID assigned to identify each conditioning variable corresponding to a particular background or subject area question within the entire pool of conditioning variables. The first four characters identify the origin of the variable: BACK (background questionnaire), SUBJ (subject specific questionnaire) SCHL (school questionnaire), TCHR (background part of teacher questionnaire), and TSUB (subject classroom part of teacher questionnaire). The second four digits represent the sequential position within each origin group.
DESCRIPTION	A short description of the conditioning variable.
GRADES/ASSESSMENTS	Three characters identifying assessment ("S" for state, "N" for national) and grade (04, 08, and 12) in which the conditioning variable was used.
CONDITIONING VAR LABEL	A descriptive eight-character label identifying the conditioning variable.
NAEP ID	The seven-character NAEP database identification for the conditioning variable.
TYPE OF CONTRAST	The type of conditioning variable. "CLASS" identifies a categorical conditioning variable and "SCALE" identifies continuous or quasi-continuous conditioning variables. "INTERACTION" identifies a set of orthogonal contrasts formed from two or more CLASS variables. "OTHER" conditioning variables do not fall into any of the above types.
TOTAL NUMBER OF SPECIFIED CONTRASTS	Each conditioning variable forms a set of one or more contrasts. For each valid response value of conditioning variable a contrast must be defined. One or more response values may be collapsed together to form one contrast. The number of response value "sets" of a conditioning variable forming a unique contrast is the value given in this field.
NUMBER OF INDEPENDENT CONTRASTS	The number of degrees of freedom in a set of contrasts formed from a conditioning variable. For a categorical conditioning variable, this number would be the number of response options minus 1 if each response option formed its own unique contrast.

Table C-2

Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
BKSER	BACK0001	GRAND MEAN (BOOK COVER)
DSEX	BACK0002	DERIVED SEX (WESTAT)
DRACE	BACK0003	DERIVED RACE (WESTAT)
B003101	BACK0004	IF HISPANIC, WHAT IS YOUR HISPANIC BACKGROUND
TOL8	BACK0005	TYPE OF LOCATION (WESTAT) (WESTAT)
TOL5	BACK0006	TYPE OF LOCATION (COLLAPSED) (WESTAT)
DOC	BACK0007	DESCRIPTION OF COMMUNITY (WESTAT)
PARED	BACK0008	PARENTS’ EDUCATION (ETS)
SCHTYPE	BACK0010	SCHOOL TYPE (PQ)
IEP	BACK0011	INDIVIDUALIZED EDUCATION PLAN (BOOK COVER)
LEP	BACK0012	LIMITED ENGLISH PROFICIENCY (BOOK COVER)
CHAP1	BACK0013	
SLUNCH	BACK0014	SCHOOL LUNCH (WESTAT)
B003201	BACK0015	HOW OFTEN OTHER THAN ENGLISH SPOKEN IN HOME
B009001	BACK0016	HOW MUCH TV/VIDEO TAPES WATCH ON A SCHOOL DAY?
B009001	BACK0017	HOW MUCH TV/VIDEO TAPES WATCH ON A SCHOOL DAY?
B006601	BACK0018	TIME SPENT ON HOMEWORK
B006601	BACK0019	TIME SPENT ON HOMEWORK
B006601	BACK0020	TIME SPENT ON HOMEWORK
B009101	BACK0021	HOURS EXTRA READING/WK, NOT CONNECTED W/ SCHOOL?
HOMEEN3	BACK0022	HOME ENVIRONMENT - ARTICLES (OF 4) IN HOME (ETS)
S004001	BACK0023	HOW MANY DAYS OF SCHOOL MISSED LAST MONTH
B007602	BACK0024	HOW MANY GRADES IN THIS STATE (4)
B010101	BACK0025	SINCE 1ST GR, NOT PROMOTION, HOW OFTEN DIFF SCHLS?
B008001	BACK0026	HOW LONG LIVED IN THE UNITED STATES
SCHNORM	SCHL0001	SCHOOL-LEVEL NORMIT GAUSSIAN SCORE (STUDENT) (ETS)
SCHNORM	SCHL0002	SCHOOL-LEVEL NORMIT GAUSSIAN SCORE (STUDENT) (ETS)
B007401	BACK0031	DISCUSS STUDIES AT HOME
B009301	BACK0032	HOW OFTEN USE A HOME COMPUTER FOR SCHOOLWORK?
SCITAKE	SUBJ0001	SCIENCE CLASS TAKING THIS YEAR (ETS)
MA93FLG	BACK0070	METRO STAT AREA (MSA) 6/30/93 DEF. (WESTAT)
MONSTUD	BACK0071	ACTUAL MONITOR STATUS (STUDENT LEVEL) (WESTAT)
SUBSAMP	BACK0073	SAMPLE TYPE (S1,S2,S3) (ETS)
B008901	BACK0075	DO YOU HAVE YOUR OWN STUDY DESK OR TABLE AT HOME?
B009401	BACK0076	HOW SAFE DO YOU FEEL AT SCHOOL?
B005601	BACK0077	DOES MOTHER OR STEPMOTHER LIVE AT HOME WITH YOU
B005701	BACK0078	DOES FATHER OR STEPFATHER LIVE AT HOME WITH YOU
K811001	SUBJ0002	AGREE/DISAGREE: I LIKE SCIENCE
K811002	SUBJ0003	AGREE/DISAGREE: I AM GOOD AT SCIENCE
K811003	SUBJ0004	AGREE/DISAGREE: LEARNING SCI MOSTLY MEMORIZATION
K811004	SUBJ0005	AGREE/DISAGREE: SCI USEFUL FOR EVERYDAY PROBLEMS
K811005	SUBJ0006	AGREE/DISAGREE: IF CHOICE, WOULD NOT STUDY SCIENCE
K811006	SUBJ0007	AGREE/DISAGREE: ALL CAN DO WELL IN SCI IF THEY TRY
K811007	SUBJ0008	AGREE/DISAGREE: SCIENCE IS BORING
K811008	SUBJ0009	AGREE/DISAGREE: SCIENCE IS A HARD SUBJECT
K811101	SUBJ0010	EVER DONE HANDS-ON PROJECT WITH LIVING THINGS?

Table C-2 (continued)
Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
K811102	SUBJ0011	EVER DONE HANDS-ON PROJECT WITH ELECTRICITY?
K811103	SUBJ0012	EVER DONE HANDS-ON PROJECT WITH CHEMICALS?
K811104	SUBJ0013	EVER DONE HANDS-ON PROJECT WITH ROCKS OR MINERALS?
K811105	SUBJ0014	DONE HANDS-ON PROJ W/ MAGNIFYING GLASS/MICROSCOPE?
K811106	SUBJ0015	DONE HANDS-ON PROJ W/ THERMOMETER OR BAROMETER?
K811107	SUBJ0016	EVER DONE HANDS-ON PROJECT WITH SIMPLE MACHINES?
K811108	SUBJ0017	HAVE DONE HANDS-ON PROJECT WITH NONE OF THE ABOVE?
K811201	SUBJ0018	HOW OFTEN DO YOU STUDY SCIENCE IN SCHOOL?
K811401	SUBJ0020	DO SCI PROJECTS IN SCHOOL THAT TAKE 1 OR MORE WKS?
K811501	SUBJ0021	LAST 2 YRS, BEEN IN SCI FAIR, FESTIVAL, SCI DAY?
K811601	SUBJ0022	FOR SCI IN SCHOOL, HOW OFTEN DO YOU READ TEXTBOOK?
K811602	SUBJ0023	FOR SCI IN SCHOOL, HOW OFTEN DO YOU READ MAGS/BKS?
K811603	SUBJ0024	FOR SCI IN SCHOOL, HOW OFTEN DISCUSS SCIENCE NEWS?
K811604	SUBJ0025	FOR SCI IN SCHOOL, HOW OFTEN WORK WITH OTHERS?
K811605	SUBJ0026	FOR SCI IN SCHOOL, HOW OFTEN GIVE ORAL REPORT?
K811606	SUBJ0027	FOR SCI IN SCHOOL, HOW OFTEN GIVE WRITTEN REPORT?
K811609	SUBJ0030	FOR SCI IN SCHOOL, HOW OFTEN DO YOU USE COMPUTER?
K811610	SUBJ0031	FOR SCI IN SCHOOL, HOW OFTEN TAKE TEST OR QUIZ?
K811611	SUBJ0032	FOR SCI IN SCHOOL, HOW OFTEN DO YOU USE LIBRARY?
K811612	SUBJ0033	FOR SCI IN SCHOOL, HOW OFTEN OBSERVE/MEAS OUTSIDE?
K811701	SUBJ0034	HOW OFTEN DOES SCIENCE TEACHER TALK TO CLASS?
K811702	SUBJ0035	HOW OFTEN DOES SCIENCE TEACHER DO DEMONSTRATION?
K811703	SUBJ0036	HOW OFTEN DOES SCIENCE TEACHER SHOW VIDEO OR TV?
K811704	SUBJ0037	HOW OFTEN DOES SCIENCE TEACHER USE COMPUTER?
K811705	SUBJ0038	HOW OFTEN DOES SCI TEACHER USE CD'S/LASER DISCS?
K811801	SUBJ0039	HOW OFTEN DOES SCI CLASS GO ON A FIELD TRIP?
K811901	SUBJ0040	HOW OFTEN DOES GUEST SPEAKER COME TO SCI CLASS?
SM00101	SUBJ0041	ON SCI TEST JUST TAKEN, NO. QUES THINK GOT RIGHT?
SM00201	SUBJ0042	COMPARED W/ OTHER SCI TESTS HOW HARD WAS THIS ONE?
SM00301	SUBJ0043	COMPARED W/ OTHER SCI TESTS, HOW HARD DID YOU TRY?
SM00401	SUBJ0044	HOW IMPT TO YOU TO DO WELL ON THIS SCIENCE TEST?
SM00501	SUBJ0045	HOW OFTEN ASKED THIS YR TO WRITE LONG SCI ANSWERS?
B009701	BACK0079	DESCRIBE YOUR OVERALL GRADES SINCE 6TH GRADE
B009801	BACK0080	HOW FAR IN SCHOOL DO YOU THINK YOU WILL GO?
B009601	BACK0083	DOES YOUR STEP/MOTHER WORK AT A JOB FOR PAY?
B009602	BACK0084	DOES YOUR STEP/FATHER WORK AT A JOB FOR PAY?
K812101	SUBJ0046	DO YOU/TEACHER SAVE YOUR SCI WORK IN A PORTFOLIO?
K812201	SUBJ0047	HOW MUCH TIME WEEKLY SPENT ON SCIENCE HOMEWORK?
K811613	SUBJ0048	FOR SCI IN SCHOOL, HOW OFTEN HANDS-ON ACTIVITIES?
K811614	SUBJ0049	FOR SCI, HOW OFTEN DISCUSS HANDS-ON RESULTS?
K811615	SUBJ0050	FOR SCI, DESIGN & CARRY OUT OWN INVESTIGATION?
C031603	SCHL0009	HAS MATH BEEN IDENTIFIED AS A PRIORITY?
C031607	SCHL0010	HAS SCIENCE BEEN IDENTIFIED AS A PRIORITY?
C031610	SCHL0012	HAS ARTS BEEN IDENTIFIED AS A PRIORITY?
C031606	SCHL0013	HAS SUBJECT INTEGRATION BEEN A PRIORITY?
C035701	SCHL0014	COMPUTERS AVAILABLE ALL THE TIME IN CLASSROOM?

Table C-2 (continued)
Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
C035702	SCHL0015	COMPUTERS GROUPED IN SEPARATE LAB AND AVAILABLE?
C035703	SCHL0016	COMPUTERS AVAILABLE TO BRING TO ROOM WHEN NEEDED?
C037201	SCHL0017	SCHOOL W/ SPECIAL FOCUS ON MATH?
C037202	SCHL0018	SCHOOL W/ SPECIAL FOCUS ON SCIENCE?
C037204	SCHL0020	SCHOOL W/ SPECIAL FOCUS ON ARTS?
C037205	SCHL0021	SCHOOL W/ SPECIAL FOCUS ON OTHER?
C037206	SCHL0022	SCHOOL NOT A SPECIAL FOCUS SCHOOL?
C037301	SCHL0023	SCHOOL FOLLOW DISTRICT/STATE MATH CURRICULUM?
C037302	SCHL0024	SCHOOL FOLLOW DISTRICT/STATE SCIENCE CURRICULUM?
C037304	SCHL0026	SCHOOL FOLLOW DISTRICT/STATE ARTS CURRICULUM?
C037305	SCHL0027	SCHOOL FOLLOW DISTRICT/STATE FOR NONE OF ABOVE?
C036601	SCHL0044	WHICH BEST DESCRIBES PRIMARY WAY LIBRARY STAFFED?
C032207	SCHL0045	INVOLVE PARENTS AS AIDES IN CLASSROOM?
C032209	SCHL0046	HAVE PARENTS REVIEW/SIGN HOMEWORK?
C032210	SCHL0047	ASSIGN HOMEWORK STUDENTS DO WITH PARENTS?
C032211	SCHL0048	HAVE A PARENT VOLUNTEER PROGRAM?
C037701	SCHL0049	WHAT % OF PARENTS IN PARENT-TEACHER ORGS?
C037702	SCHL0050	WHAT % OF PARENTS IN OPEN HOUSE/BACK SCHOOL NIGHT?
C037703	SCHL0051	WHAT % OF PARENTS IN PARENT-TEACHER COMFERENCES?
C037704	SCHL0052	WHAT % PARENTS INVOLVED MAKING CURRICULUM DECISION
C037705	SCHL0053	WHAT % OF PARENTS IN VOLUNTEER PROGRAMS?
C032402	SCHL0054	IS STUDENT ABSENTEEISM A PROBLEM IN YOUR SCHOOL?
C032401	SCHL0055	IS STUDENT TARDINESS A PROBLEM IN YOUR SCHOOL?
C032404	SCHL0056	ARE PHYSICAL CONFLICTS A PROBLEM IN YOUR SCHOOL?
C032406	SCHL0057	IS TEACHER ABSENTEEISM A PROBLEM IN YOUR SCHOOL?
C032407	SCHL0058	ARE RACE/CULT. CONFLICTS A PROBLEM IN YOUR SCHOOL?
C032408	SCHL0059	IS STUDENT HEALTH A PROBLEM IN YOUR SCHOOL?
C032409	SCHL0060	IS LACK OF PARENT INVLMT A PROBLEM IN YOUR SCHOOL?
C032410	SCHL0061	IS STUD USE OF ALCOHOL A PROBLEM IN YOUR SCHOOL?
C032411	SCHL0062	IS STUDENT TOBACCO USE A PROBLEM IN YOUR SCHOOL?
C032412	SCHL0063	IS STUDENT DRUG USE A PROBLEM IN YOUR SCHOOL?
C032413	SCHL0064	ARE GANG ACTIVITIES A PROBLEM IN YOUR SCHOOL?
C032414	SCHL0065	IS STUDENT MISBEHAVIOR A PROBLEM IN YOUR SCHOOL?
C032415	SCHL0066	IS STUDENT CHEATING A PROBLEM IN YOUR SCHOOL?
C032502	SCHL0067	IS TEACHER MORALE POS. OR NEG.?
C032503	SCHL0068	ARE STUDENT ATTITUDES TO ACADEMICS POS. OR NEG.?
C032505	SCHL0069	IS PARENT SUPPORT FOR ACHIEVEMENT POS. OR NEG.?
C032506	SCHL0070	IS REGARD FOR SCHOOL PROPERTY POS. OR NEG.?
C033601	SCHL0071	% ABSENT ON AVERAGE DAY?
C036501	SCHL0072	WHAT % OF TEACHERS ABSENT ON GIVEN DAY?
C037801	SCHL0073	% OF STUDS EROLLED AT START OF YR EROLLED AT END?
C038001	SCHL0075	% OF FULL TIME TEACHERS LEFT BEFORE END OF YR?
C038301	SCHL0076	IS SCHOOL IN NATIONAL SCHOOL LUNCH PROGRAM?
C038801	SCHL0077	SCHOOL RECEIVE CHAP 1/TITLE 1 FUNDING?
C034101	SCHL0078	DID PRINCIPAL FILL OUT THIS QUESTIONNAIRE
C034102	SCHL0079	DID HEADMASTER/HEADMISTRESS FILL OUT QUESTIONNAIRE
C034103	SCHL0080	DID HEAD TEACHER FILL OUT THIS QUESTIONNAIRE

Table C-2 (continued)
Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
C034104	SCHL0081	DID VICE PRINCIPAL FILL OUT THIS QUESTIONNAIRE
C034105	SCHL0082	DID COUNSELOR FILL OUT THIS QUESTIONNAIRE
C034106	SCHL0083	DID CURRICULUM COORD FILL OUT THIS QUESTIONNAIRE
C034107	SCHL0084	DID TEACHER FILL OUT THIS QUESTIONNAIRE
C034108	SCHL0085	DID SECRETARY FILL OUT THIS QUESTIONNAIRE
C034109	SCHL0086	DID OTHER PERSON FILL OUT THIS QUESTIONNAIRE
T055901	TCHR0001	WHAT IS YOUR GENDER?
T056001	TCHR0002	WHICH BEST DESRIBES YOU?
T040301	TCHR0003	YEARS TAUGHT
T056201	TCHR0006	TYPE TCHNG CERT IN THIS ST IN MAIN FIELD?
T040501	TCHR0007	CERTIFICATION, ELEMENTARY OR MIDDLE/JUNIOR HS ED?
T040507	TCHR0010	CERTIFICATION, ELEMENTARY SCIENCE?
T040508	TCHR0011	CERTIFICATION, MIDDLE/JUNIOR SCIENCE
T040505	TCHR0012	CERTIFICATION, OTHER
T056301	TCHR0013	HIGHEST ACADEMIC DEGREE YOU HOLD?
T040701	TCHR0014	EDUCATION UNDERGRAD MAJOR
T040706	TCHR0015	ELMENT ED UNDERGRAD MAJOR
T040707	TCHR0016	SEC ED UNDERGRAD MAJOR
T040710	TCHR0019	SCIENCE ED UNDERGRAD MAJOR
T040711	TCHR0020	LIFE SCIENCE UNDERGRAD MAJOR?
T040712	TCHR0021	PHYSICAL SCIENCE UNDERGRAD MAJOR?
T040713	TCHR0022	EARTH SCIENCE UNDERGRAD MAJOR?
T040708	TCHR0023	SPECIAL EDUCATION UNDERGRAD MAJOR
T040709	TCHR0024	BILINGUAL ED/ESL UNDERGRAD MAJOR
T040705	TCHR0025	OTHER UNDERGRAD MAJOR
T040801	TCHR0026	EDUCATION GRAD MAJOR
T040807	TCHR0027	ELEMENTARY ED GRAD MAJOR
T040808	TCHR0028	SECONDARY ED GRAD MAJOR
T040814	TCHR0031	SCIENCE ED GRAD MAJOR?
T040815	TCHR0032	LIFE SCIENCE GRAD MAJOR?
T040816	TCHR0033	PHYSICAL SCIENCE GRAD MAJOR?
T040817	TCHR0034	EARTH SCIENCE GRAD MAJOR?
T040809	TCHR0035	SPECIAL ED GRAD MAJOR
T040810	TCHR0036	BILINGUAL GRAD MAJOR
T040811	TCHR0037	ADMIN/SUPERVISION GRAD MAJOR
T040812	TCHR0038	CURRICULUM/INSTRUCTION GRAD MAJOR?
T040813	TCHR0039	COUNSELING GRAD MAJOR?
T040805	TCHR0040	OTHER GRAD MAJOR
T040806	TCHR0041	NO GRADUATE STUDY
T056401	TCHR0042	UNDERGRAD/GRAD MINOR STUDY-EDUCATION
T056402	TCHR0043	UNDERGRAD/GRAD MINOR STUDY-ELEMENTARY ED
T056403	TCHR0044	UNDERGRAD/GRAD MINOR STUDY-SECONDARY ED
T056405	TCHR0046	UNDERGRAD/GRAD MINOR STUDY-MATHEMATICS ED
T056413	TCHR0047	UNDERGRAD/GRAD MINOR STUDY-SCIENCE ED
T056414	TCHR0048	UNDERGRAD/GRAD MINOR STUDY-LIFE SCIENCE
T056415	TCHR0049	UNDERGRAD/GRAD MINOR STUDY-PHYSICAL SCIENCE
T056406	TCHR0051	UNDERGRAD/GRAD MINOR STUDY-SPECIAL ED

Table C-2 (continued)
Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
T056407	TCHR0052	UNDERGRAD/GRAD MINOR STUDY-BILINGUAL ED
T056408	TCHR0053	UNDERGRAD/GRAD MINOR STUDY-ADMIN & SUPERVISION
T056409	TCHR0054	UNDERGRAD/GRAD MINOR STUDY-CURRICULUM & INSTRUC
T056410	TCHR0055	UNDERGRAD/GRAD MINOR STUDY-COUNSELING
T056411	TCHR0056	UNDERGRAD/GRAD MINOR STUDY-OTHER
T056412	TCHR0057	UNDERGRAD/GRAD MINOR STUDY-NONE
T056701	TCHR0062	PAST 5 YRS, TAKEN COURSES/IN PRO DEVP-TELECOMM USE
T056702	TCHR0063	PAST 5 YRS, TAKEN COURSES/IN PRO DEVP-TECH USE
T056703	TCHR0064	PAST 5 YRS, TAKEN COURSES/IN PRO DEVP-COOP INSTRCT
T056704	TCHR0065	PAST 5 YRS, COURSES/IN PRO DEVLP-INTERDISP INSTRCT
T056705	TCHR0066	PAST 5 YRS, COURSES/IN PRO DEVLP-PORTFOLIO ASSMNT
T056706	TCHR0067	PAST 5 YRS, COURSES/IN PRO DEVLP-PERF BASED ASSMNT
T056707	TCHR0068	PAST 5 YRS, COURSES/PRO DEVLP-TEACH HIGHORDER THKG
T056708	TCHR0069	PAST 5 YRS, COURSES/PRO DEVLP-TEACH DIFF CULT BKGD
T056709	TCHR0070	PAST 5 YRS, COURSES/PRO DEVLP-TEACH LEP STUDENTS
T056710	TCHR0071	PAST 5 YRS, COURSES/PRO DEVLP-TEACH SPEC NEED STDS
T056711	TCHR0072	PAST 5 YRS, COURSES/PRO DEVLP-CLASSRM MNGMT/ORG
T056712	TCHR0073	PAST 5 YRS, COURSES/PRO DEVLP-OTHER PROF ISSUES
T056713	TCHR0074	PAST 5 YRS, COURSES/PRO DEVLP-NONE OF ABOVE
T041201	TCHR0075	AVAILABILITY OF RESOURCES
T056801	TCHR0078	HOW MANY SCHOOL HOURS ARE PREP TIME PER WEEK?
T060301	TCHR0080	METHODS OF TCHING SCI? COLLEGE COURSE
T060311	TCHR0081	METHODS OF TCHING SCI?WRKSHP >1 WK
T060321	TCHR0082	METHODS OF TCHING SCI?WRKSHP <1 WK >1 DAY
T060331	TCHR0083	METHODS OF TCHING SCI?WRKSHP <= 1 DAY
T060341	TCHR0084	METHODS OF TCHING SCI?OTHER PROF. DEV
T060302	TCHR0085	UNIV COURSES IN-BIO/LIFE SCI? COLLEGE COURSE
T060312	TCHR0086	UNIV COURSES IN-BIO/LIFE SCI?WRKSHP >1 WK
T060322	TCHR0087	UNIV COURSES IN-BIO/LIFE SCI?WRKSHP <1 WK >1 DAY
T060332	TCHR0088	UNIV COURSES IN-BIO/LIFE SCI?WRKSHP <= 1 DAY
T060342	TCHR0089	UNIV COURSES IN-BIO/LIFE SCI?OTHER PROF. DEV
T060303	TCHR0090	UNIV COURSES IN-CHEMISTRY? COLLEGE COURSE
T060313	TCHR0091	UNIV COURSES IN-CHEMISTRY?WRKSHP >1 WK
T060323	TCHR0092	UNIV COURSES IN-CHEMISTRY?WRKSHP <1 WK >1 DAY
T060333	TCHR0093	UNIV COURSES IN-CHEMISTRY?WRKSHP <= 1 DAY
T060343	TCHR0094	UNIV COURSES IN-CHEMISTRY?OTHER PROF. DEV
T060304	TCHR0095	UNIV COURSES IN-PHYSICS? COLLEGE COURSE
T060314	TCHR0096	UNIV COURSES IN-PHYSICS?WRKSHP >1 WK
T060324	TCHR0097	UNIV COURSES IN-PHYSICS?WRKSHP <1 WK >1 DAY
T060334	TCHR0098	UNIV COURSES IN-PHYSICS?WRKSHP <= 1 DAY
T060344	TCHR0099	UNIV COURSES IN-PHYSICS?OTHER PROF. DEV
T060305	TCHR0100	UNIV COURSES IN-EARTH SCI? COLLEGE COURSE
T060315	TCHR0101	UNIV COURSES IN-EARTH SCI?WRKSHP >1 WK
T060325	TCHR0102	UNIV COURSES IN-EARTH SCI?WRKSHP <1 WK >1 DAY
T060335	TCHR0103	UNIV COURSES IN-EARTH SCI?WRKSHP <= 1 DAY
T060345	TCHR0104	UNIV COURSES IN-EARTH SCI?OTHER PROF. DEV
T060306	TCHR0105	UNIV COURSES-OTHER TYPES OF SCI? COLLEGE COURSE

Table C-2 (continued)
Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
T060316	TCHR0106	UNIV COURSES-OTHER TYPES OF SCI?WRKSHP >1 WK
T060326	TCHR0107	UNIV COURSES-OTHR TYPES OF SCI?WRKSHP <1 WK >1 DAY
T060336	TCHR0108	UNIV COURSES-OTHER TYPES OF SCI?WRKSHP <= 1 DAY
T060346	TCHR0109	UNIV COURSES-OTHER TYPES OF SCI?OTHER PROF. DEV
T060307	TCHR0110	UNIV COURSES IN-NONE OF ABOVE? COLLEGE COURSE
T060317	TCHR0111	UNIV COURSES IN-NONE OF ABOVE?WRKSHP >1 WK
T060327	TCHR0112	UNIV COURSES IN-NONE OF ABOVE?WRKSHP <1 WK >1 DAY
T060337	TCHR0113	UNIV COURSES IN-NONE OF ABOVE?WRKSHP <= 1 DAY
T060347	TCHR0114	UNIV COURSES IN-NONE OF ABOVE?OTHER PROF. DEV
T060401	TCHR0115	PAST 5 YRS, COURSES/ACTVTS IN-COMP USE TO GET DATA
T060402	TCHR0116	PAST 5 YRS, COURSES/ACTVTS IN-COMP DATA ANALYSIS?
T060403	TCHR0117	PAST 5 YRS, COURSES/ACTVTS IN-MULTIMEDIA SCI ED?
T060404	TCHR0118	PAST 5 YRS, COURSES/ACTVTS IN-LAB MNGMT/SAFETY?
T060405	TCHR0119	PAST 5 YRS, COURSES/ACTVTS IN-INTEGRATED SCI INST?
T060501	TCHR0120	YOU BELONG TO 1 OR > SCI RELATED SCI ORGS?
T060601	TSUB0001	HOW OFTEN STUDS READ SCI TEXTBOOK?
T060602	TSUB0002	HOW OFTEN STUDS READ BOOK/MAN ABOUT SCI?
T060603	TSUB0003	HOW OFTEN STUDS DISCUSS SCI IN THE NEWS?
T060604	TSUB0004	HOW OFTEN STUDS WORK W/ OTHER STUDS ON ACT/PROJCT?
T060605	TSUB0005	HOW OFTEN STUDS GIVE ORAL SCI REPORT?
T060606	TSUB0006	HOW OFTEN STUDS PREPARE A WRITTEN SCI REPORT?
T060607	TSUB0007	HOW OFTEN STUDS DO HANDS ON SCI ACTIVITIES IN SCI?
T060608	TSUB0008	HOW OFTEN STUDS TALK ABOUT MEASURES/RESULTS?
T060609	TSUB0009	HOW OFTEN STUDS TAKE SCI TEST OR QUIZ?
T060610	TSUB0010	HOW OFTEN STUDS USE LIBRARY RESOURCES FOR SCI?
T060611	TSUB0011	HOW OFTEN STUDS USE COMPUTERS FOR SCI?
T060701	TSUB0012	HOW OFTEN DO YOU TALK TO CLASS ABOUT SCI?
T060702	TSUB0013	HOW OFTEN DO YOU DO A SCI DEMONSTRATION?
T060703	TSUB0014	HOW OFTEN DO YOU SHOW A SCI VIDEOTAPE/TV PROGRAM?
T060704	TSUB0015	HOW OFTEN DO YOU USE COMPUTERS FOR SCI?
T060705	TSUB0016	HOW OFTEN DO YOU USE CDS OR LASER DISKS ON SCI?
T060801	TSUB0017	HOW OFTEN YOUR SCI STUDS GO ON SCI FIELD TRIPS?
T060901	TSUB0018	HOW OFTEN DO YOU BRING GUEST SPEAKER FOR SCI STUDS
T061001	TSUB0019	SAVE STUDS SCI WORK IN PORTFOLIOS FOR ASSESSMENT?
T061101	TSUB0020	HOW MUCH EMPHASIS-KNOWING SCI FACTS/TERMS?
T061102	TSUB0021	HOW MUCH EMPHASIS-UNDERSTANDING KEY SCI CONCEPTS?
T061103	TSUB0022	HOW MUCH EMPHASIS-DEVELOP SCI PROB SOLVING SKILL?
T061104	TSUB0023	HOW MUCH EMPHASIS-SCI RELEVANCE TO SOCIETY/TECH?
T061105	TSUB0024	HOW MUCH EMPHASIS-COMMUNICATE IDEAS IN SCI?
T061106	TSUB0025	HOW MUCH EMPHASIS-DEVELOPING LAB SKILLS?
T061107	TSUB0026	HOW MUCH EMPHASIS-DEVELOPING STUDS SCI INTEREST?
T061108	TSUB0027	HOW MUCH EMPHASIS-DEVELOPING DATA ANALYSIS SKILLS
T061109	TSUB0028	HOW MUCH EMPHASIS-USING TECH AS SCI TOOL?
T061201	TSUB0029	EVER ASSIGN SOLO/GROUP SCI PROJECTS THAT TAKE >WK?
T061301	TSUB0030	HOW OFTEN USE MULT CHOICE TESTS TO ACCESS?
T061302	TSUB0031	HOW OFTEN USE SHOR/LONG WRITTEN RESPONSE TO ACCESS
T061303	TSUB0032	HOW OFTEN USE SOLO PROJECTS TO ACCESS?

Table C-2 (continued)

Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
T061304	TSUB0033	HOW OFTEN USE GROUP PROJECTS TO ACCESS?
T061305	TSUB0034	HOW OFTEN USE WORK PORTFOLIOS TO ACCESS?
T061306	TSUB0035	HOW OFTEN USE IN CLASS ESSAYS TO ACCESS?
T061307	TSUB0036	HOW OFTEN USE SELF/PEER EVAL TO ACCESS?
T061308	TSUB0037	HOW OFTEN USE LAB NOTEBOOKS/JOURNALS TO ACCESS?
T061309	TSUB0038	HOW OFTEN USE HOMEWORK TO ACCESS?
T061310	TSUB0039	HOW OFTEN USE HANDS ON ACTIVITIES TO ACCESS?
T061401	TSUB0040	PROPORTION OF EVAL IN SCI BASED ON HANDS ON ACTVS?
T061501	TSUB0041	BEST DESCRIPTION OF COMPUTER AVAILABILITY FOR SCI
T061601	TSUB0042	USE COMPUTERS FOR SCI INSTRUCTION: PLAYING SCI
T061611	TSUB0043	USE COMPUTERS FOR SCI INSTRUCTION: PLAYING SCI
T061621	TSUB0044	USE COMPUTERS FOR SCI INSTRUCTION: SIMULATIONS
T061631	TSUB0045	USE COMPUTERS FOR SCI INSTRUCTION: DATA ANALYSIS
T061641	TSUB0046	USE COMPUTERS FOR SCI INSTRUCTION: WORD PROCESS
T061651	TSUB0047	USE COMPUTERS FOR SCI INSTRUCTION: DO NOT USE
T061701	TSUB0048	STUDS ASSIGNED TO CLASS BY ABILITY/ACHVMNT LEVEL?
T061801	TSUB0049	IF ASSIGNED BY ABILITY WHICH BEST DESCRIBES LEVEL?
T061901	TSUB0050	COMPOSITION OF CLASS ACCORDING TO GENDER?
T062001	TSUB0051	HOW MUCH TIME CLASS SPEND ON LIFE SCIENCE?
T062002	TSUB0052	HOW MUCH TIME CLASS SPEND ON EARTH SCIENCE?
T062003	TSUB0053	HOW MUCH TIME CLASS SPEND ON PHYSICAL SCIENCE?
T062101	TSUB0054	WHICH BEST DESCRIBES SPACE WHERE CLASS TAUGHT?
T062201	TSUB0055	DO STUDS PRODUCE NOTEBOOKS/REPORTS OF LAB WORK?
T062202	TSUB0056	DO STUDS PRODUCE REPORTS OF EXTENDED SCI PROJECTS?
T062203	TSUB0057	DO STUDS PRODUCE REPORTS ON SPECIFIC TOPIC/ISSUE?
T062204	TSUB0058	DO STUDS PRODUCE REPORTS/RECORDS OF FIELD TRIPS?
T062205	TSUB0059	DO STUDS PRODUCE JOURNALS/DIARIES/LOGS OF IDEAS?
T062206	TSUB0060	DO STUDS PRODUCE PHOTO RECORDS OF PROJECTS?
T062207	TSUB0061	DO STUDS PRODUCE AUDIO/VIDEOTAPE RECORDS OF ACTVS?
T062208	TSUB0062	DO STUDS PRODUCE REPORTS OF PERSONAL INTERVIEWS?
T062209	TSUB0063	DO STUDS PRODUCE 3D SCI MODELS?
T062210	TSUB0064	DO STUDS PRODUCE COMP GENERATED MULTMEDIA PROJECTS
T062301	TSUB0065	TIME PER WEEK EXPECT STUD TO SPEND ON HOMEWORK?
T062401	TSUB0066	CLASS PERIOD AND # OF STUDS IN CLASS
C034201	SCHL0087	BEST DESCRIBES HOW 8TH GRADES ARE ORGANIZED?
C034402	SCHL0088	ARE 8TH-GRADERS ASSIGNED TO MATH BY ABILITY?
C034403	SCHL0089	ARE 8TH-GRADERS ASSIGNED TO SCIENCE BY ABILITY?
C034401	SCHL0090	ARE 8TH-GRADERS ASSIGNED TO ENGLISH BY ABILITY?
C034406	SCHL0091	ARE 8TH-GRADERS ASSIGNED TO ARTS BY ABILITY?
C034510	SCHL0092	HOW OFTEN 8TH GRDS RECEIVE COMP SCI INSTRUCTION?
C034511	SCHL0093	HOW OFTEN 8TH GRDS RECEIVE MATH INSTRUCTION?
C034512	SCHL0094	HOW OFTEN 8TH GRDS RECEIVE SCIENCE INSTRUCTION?
C034513	SCHL0095	HOW OFTEN 8TH GRDS RECEIVE ENGLISH INSTRUCTION?
C034514	SCHL0096	HOW OFTEN 8TH GRDS RECEIVE ARTS INSTRUCTION?
C031611	SCHL0097	HAS ENGLISH BEEN IDENTIFIED AS A PRIORITY?
C034601	SCHL0098	SCHOOL OFFER 8TH GR STUDS ALGEBRA FOR HS CREDIT?
C037203	SCHL0099	SCHOOL W/ SPECIAL FOCUS ON ENGLISH?

Table C-2 (continued)
Summary Table of Conditioning Variable Specifications
for the 1996 State Assessment in Science

Conditioning ID	NAEP ID	Description
C037306	SCHL0100	SCHOOL FOLLOW DISTRICT/STATE ENGLISH CURRICULUM?
C039401	SCHL0101	SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR MATH?
C039402	SCHL0102	SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR SCIENCE?
C039403	SCHL0103	SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR READING?
C039404	SCHL0104	SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR ARTS?
C039405	SCHL0105	SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR OTHER?
C039406	SCHL0106	SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR NONE ABOVE?
C039501	SCHL0107	8TH GRADERS IN EXTRACURR ACTS FOR MATH?
C039502	SCHL0108	8TH GRADERS IN EXTRACURR ACTS FOR SCIENCE?
C039503	SCHL0109	8TH GRADERS IN EXTRACURR ACTS FOR ENG/LANG ARTS?
C039504	SCHL0110	8TH GRADERS IN EXTRACURR ACTS FOR ARTS?
C039505	SCHL0111	8TH GRADERS IN EXTRACURR ACTS FOR NONE OF ABOVE?
C039601	SCHL0112	8TH GRADERS IN SUMMER PROGRAMS IN MATH?
C039602	SCHL0113	8TH GRADERS IN SUMMER PROGRAMS IN SCIENCE?
C039603	SCHL0114	8TH GRADERS IN SUMMER PROGRAMS IN ENG/LANG ARTS?
C039604	SCHL0115	8TH GRADERS IN SUMMER PROGRAMS IN ARTS?
C039605	SCHL0116	8TH GRADERS IN SUMMER PROGRAMS IN NONE OF ABOVE?
C041901	SCHL0117	WHAT % OF 8TH GRDS HELD BACK/REPEAT 8TH GRADE?
T062501	TCHR0121	COUNTING THIS YR, HOW MANY YRS TOTAL TAUGHT SCI?
T062601	TCHR0122	LAST YR, TIME IN PRO WORKSHOPS/SEMS IN SCI?
T062701	TCHR0123	LAST 2 YRS, # OF UNIV COURSES IN SCI/SCI ED?
T062801	TCHR0124	CURRICULUM SPECIALIST TO HELP/ADVISE IN SCI?
NTLUNSC	SCHL0173	PERCENT IN NATIONAL LUNCH PROGRAM FOR SCHOOL (ETS)
NTLUNSC	SCHL0174	PERCENT IN NATIONAL LUNCH PROGRAM FOR SCHOOL (ETS)
REMRDSC	SCHL0175	PERCENT IN REMEDIAL READING PRGRM FOR SCHOOL (ETS)
REMRDSC	SCHL0176	PERCENT IN REMEDIAL READING PRGRM FOR SCHOOL (ETS)
REMMHSC	SCHL0177	PERCENT IN REMEDIAL MATH PROGRAM FOR SCHOOL (ETS)
REMMHSC	SCHL0178	PERCENT IN REMEDIAL MATH PROGRAM FOR SCHOOL (ETS)
NTLUNGR	SCHL0185	PERCENT IN NATIONAL LUNCH PROGRAM FOR GRADE (ETS)
NTLUNGR	SCHL0186	PERCENT IN NATIONAL LUNCH PROGRAM FOR GRADE (ETS)
REMRDGR	SCHL0187	PERCENT IN REMEDIAL READING PRGRM FOR GRADE (ETS)
REMRDGR	SCHL0188	PERCENT IN REMEDIAL READING PRGRM FOR GRADE (ETS)
REMMHGR	SCHL0189	PERCENT IN REMEDIAL MATH PROGRAM FOR GRADE (ETS)
REMMHGR	SCHL0190	PERCENT IN REMEDIAL MATH PROGRAM FOR GRADE (ETS)

```
CONDITIONING VARIABLE ID:  BACK0001
DESCRIPTION:               GRAND MEAN
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    OVERALL
NAEP ID:                   BKSER                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          OTHER                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 OVERALL  (@          ) 1                              GRAND MEAN

CONDITIONING VARIABLE ID:  BACK0002
DESCRIPTION:               DERIVED SEX
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    GENDER
NAEP ID:                   DSEX                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 MALE     (1          ) 0                              MALE
002 FEMALE   (2          ) 1                              FEMALE

CONDITIONING VARIABLE ID:  BACK0003
DESCRIPTION:               DERIVED RACE/ETHNICITY
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    RACE/ETH
NAEP ID:                   DRACE                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         3

001 WHI/AI/O (1,5,6      ) 000                            RACE/ETHNICITY:  WHITE, AMERICAN INDIAN/ALASKAN NATIVE, OTHER,
                                                           MISSING, UNCLASSIFIED
002 BLACK    (2          ) 100                            RACE/ETHNICITY:  BLACK
003 HISPANIC (3          ) 010                            RACE/ETHNICITY:  HISPANIC
004 ASIAN    (4          ) 001                            RACE/ETHNICITY:  ASIAN

CONDITIONING VARIABLE ID:  BACK0004
DESCRIPTION:               IF HISPANIC, WHAT IS YOUR HISPANIC BACKGROUND?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    HISPANIC
NAEP ID:                   B003101                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         4

001 NOT HISP (1          ) 0000                           HISPANIC:  NOT HISPANIC
002 MEXICAN  (2          ) 1000                           HISPANIC:  MEXICAN, MEXICAN AMERICAN, CHICANO
003 PUER RIC (3          ) 0100                           HISPANIC:  PUERTO RICAN
004 CUBN,OTH (4,5        ) 0010                           HISPANIC:  CUBAN, OTHER
005 HISP-?   (M          ) 0001                           HISPANIC:  MISSING

CONDITIONING VARIABLE ID:  BACK0005
DESCRIPTION:               MSA/NON-MSA
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    MSANAT
NAEP ID:                   TOL8                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 MSA      (1,2,3,4,7,9) 0                              MSA
002 NON MSA  (5,6,8      ) 1                              NON-MSA

CONDITIONING VARIABLE ID:  BACK0006
DESCRIPTION:               TYPE OF LOCALE (5 CATEGORIES)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    TOL5
NAEP ID:                   TOL5                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         4

001 BIG CTY5 (1          ) 0000                           TOL5:  LARGE CITY
002 MID CTY5 (2,M        ) 1000                           TOL5:  MID-SIZE CITY
003 FR/BTWN5 (3          ) 0100                           TOL5:  URBAN FRINGE OF LARGE CITY, URBAN FRINGE OF MID-SIZE CITY
004 SML TWN5 (4          ) 0010                           TOL5:  SMALL TOWN
005 RURAL5   (5          ) 0001                           TOL5:  RURAL (MSA AND NON-MSA)

CONDITIONING VARIABLE ID:  BACK0007
DESCRIPTION:               DESCRIPTION OF COMMUNITY
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    DOC
NAEP ID:                   DOC                            TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         3

001 BIG CITY (1          ) 000                            DOC:  BIG CITY
002 URBAN FR (2          ) 100                            DOC:  URBAN FRINGE
003 MED CITY (3,9,M      ) 010                            DOC:  MEDIUM CITY
004 SM PLACE (4          ) 001                            DOC:  SMALL PLACE

CONDITIONING VARIABLE ID:  BACK0008
DESCRIPTION:               PARENTS' HIGHEST LEVEL OF EDUCATION
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    PARED
NAEP ID:                   PARED                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         4
```

```
001 < HS     (1          ) 0000                           PARED:  LESS THAN HIGH SCHOOL
002 HS GRAD  (2          ) 1000                           PARED:  HIGH SCHOOL GRADUATE
003 POST HS  (3          ) 0100                           PARED:  POST HIGH SCHOOL
004 COL GRAD (4          ) 0010                           PARED:  COLLEGE GRADUATE
005 PARED-?  (5,M        ) 0001                           PARED:  MISSING, I DON'T KNOW

CONDITIONING VARIABLE ID:  BACK0010
DESCRIPTION:               SCHOOL TYPE (PQ)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    SCHTYPE
NAEP ID:                   SCHTYPE                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         2

001 PUBLIC   (1          ) 00                             SCHOOL TYPE:  PUBLIC
002 PRIVATE  (2,4,5,M    ) 10                             SCHOOL TYPE:  PRIVATE, BUREAU OF INDIAN AFFAIRS, DEPARTMENT OF
                                                          DEFENSE, MISSING
003 CATHOLIC (3          ) 01                             SCHOOL TYPE:  CATHOLIC

CONDITIONING VARIABLE ID:  BACK0011
DESCRIPTION:               INDIVIDUALIZED EDUCATION PLAN
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    IEP
NAEP ID:                   IEP                            TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 IEP-YES  (1          ) 0                              IEP:  YES
002 IEP-NO   (2          ) 1                              IEP:  NO

CONDITIONING VARIABLE ID:  BACK0012
DESCRIPTION:               LIMITED ENGLISH PROFICIENCY
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    LEP
NAEP ID:                   LEP                            TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 LEP-YES  (1          ) 0                              LEP:  YES
002 LEP-NO   (2          ) 1                              LEP:  NO

CONDITIONING VARIABLE ID:  BACK0013
DESCRIPTION:               CHAPTER 1 (BOOK COVER)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    CHAPTER1
NAEP ID:                   CHAP1                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 CHAP1-Y  (1          ) 0                              CHAPTER 1:  YES
002 CHAP1-N  (2          ) 1                              CHAPTER 1:  NO

CONDITIONING VARIABLE ID:  BACK0014
DESCRIPTION:               DO YOU RECEIVE A FREE OR REDUCED-PRICE LUNCH?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    LUNCH
NAEP ID:                   SLUNCH                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         4

001 NOT ELIG (1          ) 0000                           LUNCH PROGRAM:  NOT ELIGIBLE
002 REDUCED  (2          ) 1000                           LUNCH PROGRAM:  REDUCED PRICE
003 FREE     (3          ) 0100                           LUNCH PROGRAM:  FREE
004 INFO NA  (4,M        ) 0010                           LUNCH PROGRAM:  INFO NOT AVAILABLE
005 SCH REF  (5          ) 0001                           LUNCH PROGRAM:  SCHOOL REFUAL

CONDITIONING VARIABLE ID:  BACK0015
DESCRIPTION:               HOW OFTEN DO THE PEOPLE IN YOUR HOME SPEAK A LANGUAGE OTHER THAN ENGLISH?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    HOMELANG
NAEP ID:                   B003201                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         3

001 HL-NEVER (1          ) 000                            HOMELANG:  NEVER
002 HL-SOME  (2          ) 100                            HOMELANG:  SOMETIMES
003 HL-ALWAY (3          ) 010                            HOMELANG:  ALWAYS
004 HL-?     (M          ) 001                            HOMELANG:  MISSING

CONDITIONING VARIABLE ID:  BACK0016
DESCRIPTION:               HOW MUCH TV/VIDEO DO YOU USUALLY WATCH EACH DAY? (LINEAR)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    TVWATCHL
NAEP ID:                   B009001                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     7
TYPE OF CONTRAST:          LINEAR                         NUMBER OF INDEPENDENT CONTRASTS:         1

001 TVLIN-0  (1          ) 0                              TV WATCHING (LINEAR) (0 TO 6+ HOURS PER DAY)
002 TVLIN-1  (2          ) 1                              TV WATCHING (LINEAR)
003 TVLIN-2  (3          ) 2                              TV WATCHING (LINEAR)
004 TVLIN-3  (4,M        ) 3                              TV WATCHING (LINEAR)
005 TVLIN-4  (5          ) 4                              TV WATCHING (LINEAR)
006 TVLIN-5  (6          ) 5                              TV WATCHING (LINEAR)
007 TVLIN-6  (7          ) 6                              TV WATCHING (LINEAR)
```

```
CONDITIONING VARIABLE ID:  BACK0017
DESCRIPTION:               HOW MUCH TV/VIDEO DO YOU USUALLY WATCH EACH DAY? (QUADRATIC)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    TVWATCHQ
NAEP ID:                   B009001                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          QUADRATIC                       NUMBER OF INDEPENDENT CONTRASTS:         1

001 TV-QUAD  (1-7,M=4    )  1.0 + -2.0*X +  1.0*X**2       TV WATCHING (QUADRATIC)

CONDITIONING VARIABLE ID:  BACK0018
DESCRIPTION:               HOMEWORK ASSIGNED?:  BASED ON TIME SPENT ON HOMEWORK EACH DAY.
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    HWASSIGN
NAEP ID:                   B006601                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         2

001 HW-MISS  (M          ) 00                              HOMEWORK ASSIGNED?:  MISSING
002 HW-NO    (1          ) 10                              HOMEWORK ASSIGNED?:  NO
003 HW-YES   (2-5        ) 01                              HOMEWORK ASSIGNED?:  YES

CONDITIONING VARIABLE ID:  BACK0019
DESCRIPTION:               HOW MUCH TIME DO YOU USUALLY SPEND ON HOMEWORK EACH DAY? (LINEAR)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    HOMEWRKL
NAEP ID:                   B006601                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          LINEAR                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 HWLIN-0  (1,2,M      ) 0                               HOMEWORK (LINEAR):  DON'T HAVE ANY, DON'T DO ANY, MISSING
002 HWLIN-1  (3          ) 1                               HOMEWORK (LINEAR):  1/2 HOUR OR LESS
003 HWLIN-2  (4          ) 2                               HOMEWORK (LINEAR):  1 HOUR
004 HWLIN-3  (5          ) 3                               HOMEWORK (LINEAR):  MORE THAN 1 HOUR

CONDITIONING VARIABLE ID:  BACK0020
DESCRIPTION:               HOW MUCH TIME DO YOU USUALLY SPEND ON HOMEWORK EACH DAY (QUADRATIC)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    HOMEWRKQ
NAEP ID:                   B006601                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          SCALE                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 HWQUAD-0 (1,2,M      ) 0                               HOMEWORK (QUADRATIC):  DON'T HAVE ANY, DON'T DO ANY, MISSING
002 HWQUAD-1 (3          ) 1                               HOMEWORK (QUADRATIC):  1/2 HOUR OR LESS
003 HWQUAD-2 (4          ) 4                               HOMEWORK (QUADRATIC):  1 HOUR
004 HWQUAD-3 (5          ) 9                               HOMEWORK (QUADRATIC):  MORE THAN 1 HOUR

CONDITIONING VARIABLE ID:  BACK0021
DESCRIPTION:               HOURS EXTRA READING/WK, NOT CONNECTED W/ SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   B009101                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     8
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         7

001 NONE     (01         ) 0000000                         NONE
002 1-2 HRS  (02         ) 1000000                         1-2 HOURS
003 3-4 HRS  (03         ) 0100000                         3-4 HOURS
004 5-6 HRS  (04         ) 0010000                         5-6 HOURS
005 7-8 HRS  (05         ) 0001000                         7-8 HOURS
006 9-10 HRS (06         ) 0000100                         9-10 HOURS
007 > 10 HRS (07         ) 0000010                         MORE THAN 10 HOURS
008 B009101M (M          ) 0000001                         MISSING

CONDITIONING VARIABLE ID:  BACK0022
DESCRIPTION:               NUMBER OF ITEMS IN THE HOME (NEWSPAPER, > 25 BOOKS, ENCYCLOPEDIA, MAGAZINES) (DERIVED)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    HOMEITMS
NAEP ID:                   HOMEEN3                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         2

001 HITEM<=2 (1,M        ) 00                              ITEMS IN HOME:  ZERO TO TWO ITEMS, MISSING
002 HITEM=3  (2          ) 10                              ITEMS IN HOME:  THREE ITEMS
003 HITEM=4  (3          ) 01                              ITEMS IN HOME:  FOUR ITEMS

CONDITIONING VARIABLE ID:  BACK0023
DESCRIPTION:               HOW MANY DAYS OF SCHOOL MISSED LAST MONTH?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    SCH MISS
NAEP ID:                   S004001                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 MISS->2  (3,4,5,M    ) 0                               DAYS OF SCHOOL MISSED:  3-4, 5-10, 10 OR MORE DAYS, MISSING
002 MISS-2<  (1,2        ) 1                               DAYS OF SCHOOL MISSED:  0-1, 2 DAYS

CONDITIONING VARIABLE ID:  BACK0024
DESCRIPTION:               HOW MANY GRADES IN THIS STATE (4)
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   B007602                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
```

```
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 < ONE GR (01         ) 0000                             LESS THAN ON GRADE
002 1-2 GRDS (02         ) 1000                             1-2 GRADES
003 3-5 GRDS (03         ) 0100                             3 -5 GRADES
004 > 5 GRDS (04         ) 0010                             5 OR MORE GRADESEVER
005 MISSING  (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  BACK0025
DESCRIPTION:               SINCE 1ST GR, NOT PROMOTION, HOW OFTEN DIFF SCHLS?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   B010101                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    9
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        8

001 B010101A (01         ) 00000000                         NONE
002 B010101B (02         ) 10000000                         1
003 B010101C (03         ) 01000000                         2
004 B010101D (04         ) 00100000                         3
005 B010101E (05         ) 00010000                         4
006 B010101F (06         ) 00001000                         5
007 B010101G (07         ) 00000100                         6
008 B010101H (08         ) 00000010                         7 OR MOREMORE
009 B010101M (M          ) 00000001                         MISSING

CONDITIONING VARIABLE ID:  BACK0026
DESCRIPTION:               HOW LONG LIVED IN THE UNITED STATES?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    YRSINUSA
NAEP ID:                   B008001                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 AALL LIF (1          ) 0000                             ALL MY LIFE
002 USA >5   (2          ) 1000                             LIVED IN US MORE THAN 5 YEARS
003 USA 3-5  (3          ) 0100                             LIVED IN US 3-5 YEARS
004 USA <3   (4          ) 0010                             LIVED IN US LESS THAN 3 YEARS
005 USA-?    (M          ) 0001                             LIVED IN US MISSING

CONDITIONING VARIABLE ID:  SCHL0001
DESCRIPTION:               SCHOOL LEVEL AVERAGE SCIENCE NORMIT (MISSING VS NON-MISSING)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    SCH NORM
NAEP ID:                   SCHNORM                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 SCHNRM-? (M          ) 0                                SCHOOL LEVEL AVERAGE SCIENCE NORMIT MISSING
002 SCHNRM-Y (@          ) 1                                SCHOOL LEVEL AVERAGE SCIENCE NORMIT NOT-MISSING

CONDITIONING VARIABLE ID:  SCHL0002
DESCRIPTION:               SCHOOL LEVEL AVERAGE SCIENCE NORMIT
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    SNRM-LIN
NAEP ID:                   SCHNORM                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          SCALE                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 SNRM-LIN (#          ) (F8.4)                           SCHOOL LEVEL AVERAGE SCIENCE NORMIT MEAN
002 SNRM-LIN (M          ) 0                                SCHOOL LEVEL AVERAGE SCIENCE NORMIT MISSING

CONDITIONING VARIABLE ID:  BACK0031
DESCRIPTION:               HOW OFTEN DO YOU DISCUSS THINGS STUDIED IN SCHOOL WITH SOMEONE AT HOME?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    DISC@HOM
NAEP ID:                   B007401                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 DIS@HOM1 (1          ) 000                              DISCUSS STUDIES AT HOME:  ALMOST EVERY DAY
002 DIS@HOM2 (2          ) 100                              DISCUSS STUDIES AT HOME:  ONCE OR TWICE A WEEK
003 DIS@HOM3 (3          ) 010                              DISCUSS STUDIES AT HOME:  ONCE OR TWICE A MONTH
004 DIS@HOM4 (4,M        ) 001                              DISCUSS STUDIES AT HOME:  NEVER OR HARDLY EVER, MISSING

CONDITIONING VARIABLE ID:  BACK0032
DESCRIPTION:               HOW OFTEN USE A HOME COMPUTER FOR SCHOOLWORK?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   B009301                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 DAILY    (01         ) 00000                            ALMOST EVERY DAY
002 1-2 WEEK (02         ) 10000                            ONCE OR TWICE A WEEK
003 1-2 MTH  (03         ) 01000                            ONCE OR TWICE A MTH
004 NEVER    (04         ) 00100                            NEVER OR HARDLY EVER
005 NO COMP  (05         ) 00010                            NO COMPUTER AT HOME
006 MISSING  (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  SUBJ0001
DESCRIPTION:               WHAT KIND OF SCI CLASS ARE YOU TAKING THIS YEAR?
GRADES/ASSESSMENTS:        N08, S08
```

```
CONDITIONING VAR LABEL:
NAEP ID:                  SCITAKE                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     7
TYPE OF CONTRAST:         CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         6

001 NO SCIEN (01         ) 000000                        NOT TAKING THIS YEAR
002 LIFESCI  (02         ) 100000                        LIFE SCIENCE
003 PHYSSCI  (03         ) 010000                        PHYSICAL SCIENCE
004 EATHSCI  (04         ) 001000                        EARTH SCIENCE
005 GEN SCI  (05         ) 000100                        GENERAL SCIENCE
006 INTESCI  (06         ) 000010                        INTEGRATED SCIENCE
007 MISSING  (M          ) 000001                        MISSING

CONDITIONING VARIABLE ID:  BACK0050
DESCRIPTION:               INTERACTION:  GENDER BY RACE/ETHNICITY
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    GEND/RAC
NAEP ID:                   N/A                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     8
TYPE OF CONTRAST:          INTERACTION                   NUMBER OF INDEPENDENT CONTRASTS:         3

001 G/R 11   (11         ) 010101                        GEND/RAC INTACT: 1. MALE     1. WHI/AI/O
002 G/R 12   (12         ) -10000                        GEND/RAC INTACT: 1. MALE     2. BLACK
003 G/R 13   (13         ) 00-100                        GEND/RAC INTACT: 1. MALE     3. HISPANIC
004 G/R 14   (14         ) 0000-1                        GEND/RAC INTACT: 1. MALE     4. ASIAN
005 G/R 21   (21         ) -1-1-1                        GEND/RAC INTACT: 2. FEMALE   1. WHI/AI/O
006 G/R 22   (22         ) 010000                        GEND/RAC INTACT: 2. FEMALE   2. BLACK
007 G/R 23   (23         ) 000100                        GEND/RAC INTACT: 2. FEMALE   3. HISPANIC
008 G/R 24   (24         ) 000001                        GEND/RAC INTACT: 2. FEMALE   4. ASIAN

CONDITIONING VARIABLE ID:  BACK0051
DESCRIPTION:               INTERACTION:  GENDER BY TYPE OF LOCALE (5 CATEGORIES)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    GEND/TOL
NAEP ID:                   N/A                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    10
TYPE OF CONTRAST:          INTERACTION                   NUMBER OF INDEPENDENT CONTRASTS:         4

001 G/T 11   (11         ) 01010101                      GEND/TOL INTACT: 1. MALE     1. BIG CTY5
002 G/T 12   (12         ) -1000000                      GEND/TOL INTACT: 1. MALE     2. MID CTY5
003 G/T 13   (13         ) 00-10000                      GEND/TOL INTACT: 1. MALE     3. FR/BTWN5
004 G/T 14   (14         ) 0000-100                      GEND/TOL INTACT: 1. MALE     4. SML TWN5
005 G/T 15   (15         ) 000000-1                      GEND/TOL INTACT: 1. MALE     5. RURAL5
006 G/T 21   (21         ) -1-1-1-1                      GEND/TOL INTACT: 2. FEMALE   1. BIG CTY5
007 G/T 22   (22         ) 01000000                      GEND/TOL INTACT: 2. FEMALE   2. MID CTY5
008 G/T 23   (23         ) 00010000                      GEND/TOL INTACT: 2. FEMALE   3. FR/BTWN5
009 G/T 24   (24         ) 00000100                      GEND/TOL INTACT: 2. FEMALE   4. SML TWN5
010 G/T 25   (25         ) 00000001                      GEND/TOL INTACT: 2. FEMALE   5. RURAL5

CONDITIONING VARIABLE ID:  BACK0052
DESCRIPTION:               INTERACTION:  GENDER BY PARENTS’ EDUCATION
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    GEND/PAR
NAEP ID:                   N/A                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    10
TYPE OF CONTRAST:          INTERACTION                   NUMBER OF INDEPENDENT CONTRASTS:         4

001 G/P 11   (11         ) 01010101                      GEND/PAR INTACT: 1. MALE     1. < HS
002 G/P 12   (12         ) -1000000                      GEND/PAR INTACT: 1. MALE     2. HS GRAD
003 G/P 13   (13         ) 00-10000                      GEND/PAR INTACT: 1. MALE     3. POST HS
004 G/P 14   (14         ) 0000-100                      GEND/PAR INTACT: 1. MALE     4. COL GRAD
005 G/P 15   (15         ) 000000-1                      GEND/PAR INTACT: 1. MALE     5. PARED-?
006 G/P 21   (21         ) -1-1-1-1                      GEND/PAR INTACT: 2. FEMALE   1. < HS
007 G/P 22   (22         ) 01000000                      GEND/PAR INTACT: 2. FEMALE   2. HS GRAD
008 G/P 23   (23         ) 00010000                      GEND/PAR INTACT: 2. FEMALE   3. POST HS
009 G/P 24   (24         ) 00000100                      GEND/PAR INTACT: 2. FEMALE   4. COL GRAD
010 G/P 25   (25         ) 00000001                      GEND/PAR INTACT: 2. FEMALE   5. PARED-?

CONDITIONING VARIABLE ID:  BACK0053
DESCRIPTION:               INTERACTION:  GENDER BY SCHOOL TYPE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    GEND/SCH
NAEP ID:                   N/A                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     6
TYPE OF CONTRAST:          INTERACTION                   NUMBER OF INDEPENDENT CONTRASTS:         2

001 G/S 11   (11         ) 0101                          GEND/SCH INTACT: 1. MALE     1. PUBLIC
002 G/S 12   (12         ) -100                          GEND/SCH INTACT: 1. MALE     2. PRIVATE
003 G/S 13   (13         ) 00-1                          GEND/SCH INTACT: 1. MALE     3. CATHOLIC
004 G/S 21   (21         ) -1-1                          GEND/SCH INTACT: 2. FEMALE   1. PUBLIC
005 G/S 22   (22         ) 0100                          GEND/SCH INTACT: 2. FEMALE   2. PRIVATE
006 G/S 23   (23         ) 0001                          GEND/SCH INTACT: 2. FEMALE   3. CATHOLIC

CONDITIONING VARIABLE ID:  BACK0054
DESCRIPTION:               INTERACTION:  RACE/ETHNICITY BY TYPE OF LOCALE (5 CATEGORIES)
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    RACE/TOL
NAEP ID:                   N/A                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    20
TYPE OF CONTRAST:          INTERACTION                   NUMBER OF INDEPENDENT CONTRASTS:        12

001 R/T 11   (11         ) 010101010101010101010101      RACE/TOL INTACT: 1. WHI/AI/O 1. BIG CTY5
002 R/T 12   (12         ) -1000000-1000000-1000000      RACE/TOL INTACT: 1. WHI/AI/O 2. MID CTY5
```

```
003 R/T 13   (13          ) 00-1000000-1000000-10000         RACE/TOL INTACT: 1. WHI/AI/O 3. FR/BTWN5
004 R/T 14   (14          ) 0000-1000000-1000000-100         RACE/TOL INTACT: 1. WHI/AI/O 4. SML TWN5
005 R/T 15   (15          ) 000000-1000000-1000000-1         RACE/TOL INTACT: 1. WHI/AI/O 5. RURAL5
006 R/T 21   (21          ) -1-1-1-10000000000000000         RACE/TOL INTACT: 2. BLACK    1. BIG CTY5
007 R/T 22   (22          ) 010000000000000000000000         RACE/TOL INTACT: 2. BLACK    2. MID CTY5
008 R/T 23   (23          ) 000100000000000000000000         RACE/TOL INTACT: 2. BLACK    3. FR/BTWN5
009 R/T 24   (24          ) 000001000000000000000000         RACE/TOL INTACT: 2. BLACK    4. SML TWN5
010 R/T 25   (25          ) 000000010000000000000000         RACE/TOL INTACT: 2. BLACK    5. RURAL5
011 R/T 31   (31          ) 00000000-1-1-1-100000000         RACE/TOL INTACT: 3. HISPANIC 1. BIG CTY5
012 R/T 32   (32          ) 000000000100000000000000         RACE/TOL INTACT: 3. HISPANIC 2. MID CTY5
013 R/T 33   (33          ) 000000000001000000000000         RACE/TOL INTACT: 3. HISPANIC 3. FR/BTWN5
014 R/T 34   (34          ) 000000000000010000000000         RACE/TOL INTACT: 3. HISPANIC 4. SML TWN5
015 R/T 35   (35          ) 000000000000000100000000         RACE/TOL INTACT: 3. HISPANIC 5. RURAL5
016 R/T 41   (41          ) 0000000000000000-1-1-1-1         RACE/TOL INTACT: 4. ASIAN    1. BIG CTY5
017 R/T 42   (42          ) 000000000000000001000000         RACE/TOL INTACT: 4. ASIAN    2. MID CTY5
018 R/T 43   (43          ) 000000000000000000010000         RACE/TOL INTACT: 4. ASIAN    3. FR/BTWN5
019 R/T 44   (44          ) 000000000000000000000100         RACE/TOL INTACT: 4. ASIAN    4. SML TWN5
020 R/T 45   (45          ) 000000000000000000000001         RACE/TOL INTACT: 4. ASIAN    5. RURAL5

CONDITIONING VARIABLE ID:  BACK0055
DESCRIPTION:               INTERACTION:  RACE/ETHNICITY BY PARENTS' EDUCATION
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    RACE/PAR
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    20
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:        12

001 R/P 11   (11          ) 010101010101010101010101         RACE/PAR INTACT: 1. WHI/AI/O 1. < HS
002 R/P 12   (12          ) -1000000-1000000-1000000         RACE/PAR INTACT: 1. WHI/AI/O 2. HS GRAD
003 R/P 13   (13          ) 00-1000000-1000000-10000         RACE/PAR INTACT: 1. WHI/AI/O 3. POST HS
004 R/P 14   (14          ) 0000-1000000-1000000-100         RACE/PAR INTACT: 1. WHI/AI/O 4. COL GRAD
005 R/P 15   (15          ) 000000-1000000-1000000-1         RACE/PAR INTACT: 1. WHI/AI/O 5. PARED-?
006 R/P 21   (21          ) -1-1-1-10000000000000000         RACE/PAR INTACT: 2. BLACK    1. < HS
007 R/P 22   (22          ) 010000000000000000000000         RACE/PAR INTACT: 2. BLACK    2. HS GRAD
008 R/P 23   (23          ) 000100000000000000000000         RACE/PAR INTACT: 2. BLACK    3. POST HS
009 R/P 24   (24          ) 000001000000000000000000         RACE/PAR INTACT: 2. BLACK    4. COL GRAD
010 R/P 25   (25          ) 000000010000000000000000         RACE/PAR INTACT: 2. BLACK    5. PARED-?
011 R/P 31   (31          ) 00000000-1-1-1-100000000         RACE/PAR INTACT: 3. HISPANIC 1. < HS
012 R/P 32   (32          ) 000000000100000000000000         RACE/PAR INTACT: 3. HISPANIC 2. HS GRAD
013 R/P 33   (33          ) 000000000001000000000000         RACE/PAR INTACT: 3. HISPANIC 3. POST HS
014 R/P 34   (34          ) 000000000000010000000000         RACE/PAR INTACT: 3. HISPANIC 4. COL GRAD
015 R/P 35   (35          ) 000000000000000100000000         RACE/PAR INTACT: 3. HISPANIC 5. PARED-?
016 R/P 41   (41          ) 0000000000000000-1-1-1-1         RACE/PAR INTACT: 4. ASIAN    1. < HS
017 R/P 42   (42          ) 000000000000000001000000         RACE/PAR INTACT: 4. ASIAN    2. HS GRAD
018 R/P 43   (43          ) 000000000000000000010000         RACE/PAR INTACT: 4. ASIAN    3. POST HS
019 R/P 44   (44          ) 000000000000000000000100         RACE/PAR INTACT: 4. ASIAN    4. COL GRAD
020 R/P 45   (45          ) 000000000000000000000001         RACE/PAR INTACT: 4. ASIAN    5. PARED-?

CONDITIONING VARIABLE ID:  BACK0056
DESCRIPTION:               INTERACTION:  RACE/ETHNICITY BY SCHOOL TYPE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    RACE/SCH
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    12
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:         6

001 R/S 11   (11          ) 010101010101                     RACE/SCH INTACT: 1. WHI/AI/O 1. PUBLIC
002 R/S 12   (12          ) -100-100-100                     RACE/SCH INTACT: 1. WHI/AI/O 2. PRIVATE
003 R/S 13   (13          ) 00-100-100-1                     RACE/SCH INTACT: 1. WHI/AI/O 3. CATHOLIC
004 R/S 21   (21          ) -1-100000000                     RACE/SCH INTACT: 2. BLACK    1. PUBLIC
005 R/S 22   (22          ) 010000000000                     RACE/SCH INTACT: 2. BLACK    2. PRIVATE
006 R/S 23   (23          ) 000100000000                     RACE/SCH INTACT: 2. BLACK    3. CATHOLIC
007 R/S 31   (31          ) 0000-1-10000                     RACE/SCH INTACT: 3. HISPANIC 1. PUBLIC
008 R/S 32   (32          ) 000001000000                     RACE/SCH INTACT: 3. HISPANIC 2. PRIVATE
009 R/S 33   (33          ) 000000010000                     RACE/SCH INTACT: 3. HISPANIC 3. CATHOLIC
010 R/S 41   (41          ) 00000000-1-1                     RACE/SCH INTACT: 4. ASIAN    1. PUBLIC
011 R/S 42   (42          ) 000000000100                     RACE/SCH INTACT: 4. ASIAN    2. PRIVATE
012 R/S 43   (43          ) 000000000001                     RACE/SCH INTACT: 4. ASIAN    3. CATHOLIC

CONDITIONING VARIABLE ID:  BACK0057
DESCRIPTION:               INTERACTION:  TYPE OF LOCALE (5 CATEGORIES) BY PARENT'S EDUCATION
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    TOL5/PAR
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    25
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:        16

001 T/P 11   (11          ) 01010101010101010101010101010101 TOL5/PAR INTACT: 1. BIG CTY5 1. < HS
002 T/P 12   (12          ) -1000000-1000000-1000000-1000000 TOL5/PAR INTACT: 1. BIG CTY5 2. HS GRAD
003 T/P 13   (13          ) 00-1000000-1000000-1000000-10000 TOL5/PAR INTACT: 1. BIG CTY5 3. POST HS
004 T/P 14   (14          ) 0000-1000000-1000000-1000000-100 TOL5/PAR INTACT: 1. BIG CTY5 4. COL GRAD
005 T/P 15   (15          ) 000000-1000000-1000000-1000000-1 TOL5/PAR INTACT: 1. BIG CTY5 5. PARED-?
006 T/P 21   (21          ) -1-1-1-1000000000000000000000000 TOL5/PAR INTACT: 2. MID CTY5 1. < HS
007 T/P 22   (22          ) 01000000000000000000000000000000 TOL5/PAR INTACT: 2. MID CTY5 2. HS GRAD
008 T/P 23   (23          ) 00010000000000000000000000000000 TOL5/PAR INTACT: 2. MID CTY5 3. POST HS
009 T/P 24   (24          ) 00000100000000000000000000000000 TOL5/PAR INTACT: 2. MID CTY5 4. COL GRAD
010 T/P 25   (25          ) 00000001000000000000000000000000 TOL5/PAR INTACT: 2. MID CTY5 5. PARED-?
011 T/P 31   (31          ) 00000000-1-1-1-10000000000000000 TOL5/PAR INTACT: 3. FR/BTWN5 1. < HS
012 T/P 32   (32          ) 00000000010000000000000000000000 TOL5/PAR INTACT: 3. FR/BTWN5 2. HS GRAD
013 T/P 33   (33          ) 00000000000100000000000000000000 TOL5/PAR INTACT: 3. FR/BTWN5 3. POST HS
```

```
014 T/P 34   (34          ) 00000000000001000000000000000000 TOL5/PAR INTACT: 3. FR/BTWN5 4. COL GRAD
015 T/P 35   (35          ) 00000000000000010000000000000000 TOL5/PAR INTACT: 3. FR/BTWN5 5. PARED-?
016 T/P 41   (41          ) 0000000000000000-1-1-1-100000000 TOL5/PAR INTACT: 4. SML TWN5 1. < HS
017 T/P 42   (42          ) 00000000000000000100000000000000 TOL5/PAR INTACT: 4. SML TWN5 2. HS GRAD
018 T/P 43   (43          ) 00000000000000000001000000000000 TOL5/PAR INTACT: 4. SML TWN5 3. POST HS
019 T/P 44   (44          ) 00000000000000000000010000000000 TOL5/PAR INTACT: 4. SML TWN5 4. COL GRAD
020 T/P 45   (45          ) 00000000000000000000000100000000 TOL5/PAR INTACT: 4. SML TWN5 5. PARED-?
021 T/P 51   (51          ) 000000000000000000000000-1-1-1-1 TOL5/PAR INTACT: 5. RURAL5   1. < HS
022 T/P 52   (52          ) 00000000000000000000000001000000 TOL5/PAR INTACT: 5. RURAL5   2. HS GRAD
023 T/P 53   (53          ) 00000000000000000000000000010000 TOL5/PAR INTACT: 5. RURAL5   3. POST HS
024 T/P 54   (54          ) 00000000000000000000000000000100 TOL5/PAR INTACT: 5. RURAL5   4. COL GRAD
025 T/P 55   (55          ) 00000000000000000000000000000001 TOL5/PAR INTACT: 5. RURAL5   5. PARED-?

CONDITIONING VARIABLE ID:  BACK0058
DESCRIPTION:               INTERACTION:  TYPE OF LOCALE (5 CATEGORIES) BY SCHOOL TYPE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    TOL5/SCH
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    15
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:         8

001 T/S 11   (11          ) 0101010101010101                 TOL5/SCH INTACT: 1. BIG CTY5 1. PUBLIC
002 T/S 12   (12          ) -100-100-100-100                 TOL5/SCH INTACT: 1. BIG CTY5 2. PRIVATE
003 T/S 13   (13          ) 00-100-100-100-1                 TOL5/SCH INTACT: 1. BIG CTY5 3. CATHOLIC
004 T/S 21   (21          ) -1-1000000000000                 TOL5/SCH INTACT: 2. MID CTY5 1. PUBLIC
005 T/S 22   (22          ) 0100000000000000                 TOL5/SCH INTACT: 2. MID CTY5 2. PRIVATE
006 T/S 23   (23          ) 0001000000000000                 TOL5/SCH INTACT: 2. MID CTY5 3. CATHOLIC
007 T/S 31   (31          ) 0000-1-100000000                 TOL5/SCH INTACT: 3. FR/BTWN5 1. PUBLIC
008 T/S 32   (32          ) 0000010000000000                 TOL5/SCH INTACT: 3. FR/BTWN5 2. PRIVATE
009 T/S 33   (33          ) 0000000100000000                 TOL5/SCH INTACT: 3. FR/BTWN5 3. CATHOLIC
010 T/S 41   (41          ) 00000000-1-10000                 TOL5/SCH INTACT: 4. SML TWN5 1. PUBLIC
011 T/S 42   (42          ) 0000000001000000                 TOL5/SCH INTACT: 4. SML TWN5 2. PRIVATE
012 T/S 43   (43          ) 0000000000010000                 TOL5/SCH INTACT: 4. SML TWN5 3. CATHOLIC
013 T/S 51   (51          ) 000000000000-1-1                 TOL5/SCH INTACT: 5. RURAL5   1. PUBLIC
014 T/S 52   (52          ) 0000000000000100                 TOL5/SCH INTACT: 5. RURAL5   2. PRIVATE
015 T/S 53   (53          ) 0000000000000001                 TOL5/SCH INTACT: 5. RURAL5   3. CATHOLIC

CONDITIONING VARIABLE ID:  BACK0059
DESCRIPTION:               INTERACTION:  PARENTS' EDUCATION BY SCHOOL TYPE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    PARE/SCH
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    15
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:         8

001 P/S 11   (11          ) 0101010101010101                 PARE/SCH INTACT: 1. < HS     1. PUBLIC
002 P/S 12   (12          ) -100-100-100-100                 PARE/SCH INTACT: 1. < HS     2. PRIVATE
003 P/S 13   (13          ) 00-100-100-100-1                 PARE/SCH INTACT: 1. < HS     3. CATHOLIC
004 P/S 21   (21          ) -1-1000000000000                 PARE/SCH INTACT: 2. HS GRAD  1. PUBLIC
005 P/S 22   (22          ) 0100000000000000                 PARE/SCH INTACT: 2. HS GRAD  2. PRIVATE
006 P/S 23   (23          ) 0001000000000000                 PARE/SCH INTACT: 2. HS GRAD  3. CATHOLIC
007 P/S 31   (31          ) 0000-1-100000000                 PARE/SCH INTACT: 3. POST HS  1. PUBLIC
008 P/S 32   (32          ) 0000010000000000                 PARE/SCH INTACT: 3. POST HS  2. PRIVATE
009 P/S 33   (33          ) 0000000100000000                 PARE/SCH INTACT: 3. POST HS  3. CATHOLIC
010 P/S 41   (41          ) 00000000-1-10000                 PARE/SCH INTACT: 4. COL GRAD 1. PUBLIC
011 P/S 42   (42          ) 0000000001000000                 PARE/SCH INTACT: 4. COL GRAD 2. PRIVATE
012 P/S 43   (43          ) 0000000000010000                 PARE/SCH INTACT: 4. COL GRAD 3. CATHOLIC
013 P/S 51   (51          ) 000000000000-1-1                 PARE/SCH INTACT: 5. PARED-?  1. PUBLIC
014 P/S 52   (52          ) 0000000000000100                 PARE/SCH INTACT: 5. PARED-?  2. PRIVATE
015 P/S 53   (53          ) 0000000000000001                 PARE/SCH INTACT: 5. PARED-?  3. CATHOLIC

CONDITIONING VARIABLE ID:  BACK0060
DESCRIPTION:               INTERACTION:  GENDER BY SCIENCE COURSES TAKING THIS YEAR
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:    GEND/
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    14
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:         6

001 G/  11   (11          ) 010101010101                    GEND/    INTACT: 1. MALE     1. NO SCIEN
002 G/  12   (12          ) -10000000000                    GEND/    INTACT: 1. MALE     2. LIFESCI
003 G/  13   (13          ) 00-100000000                    GEND/    INTACT: 1. MALE     3. PHYSSCI
004 G/  14   (14          ) 0000-1000000                    GEND/    INTACT: 1. MALE     4. EATHSCI
005 G/  15   (15          ) 000000-10000                    GEND/    INTACT: 1. MALE     5. GEN SCI
006 G/  16   (16          ) 00000000-100                    GEND/    INTACT: 1. MALE     6. INTESCI
007 G/  17   (17          ) 0000000000-1                    GEND/    INTACT: 1. MALE     7. MISSING
008 G/  21   (21          ) -1-1-1-1-1-1                    GEND/    INTACT: 2. FEMALE   1. NO SCIEN
009 G/  22   (22          ) 010000000000                    GEND/    INTACT: 2. FEMALE   2. LIFESCI
010 G/  23   (23          ) 000100000000                    GEND/    INTACT: 2. FEMALE   3. PHYSSCI
011 G/  24   (24          ) 000001000000                    GEND/    INTACT: 2. FEMALE   4. EATHSCI
012 G/  25   (25          ) 000000010000                    GEND/    INTACT: 2. FEMALE   5. GEN SCI
013 G/  26   (26          ) 000000000100                    GEND/    INTACT: 2. FEMALE   6. INTESCI
014 G/  27   (27          ) 000000000001                    GEND/    INTACT: 2. FEMALE   7. MISSING

CONDITIONING VARIABLE ID:  BACK0062
DESCRIPTION:               INTERACTION:  RACE/ETHNICITY BY SCIENCE COURSES TAKING THIS YEAR
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:    RACE/
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    28
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:        18
```

```
001 R/  11   (11         ) 333333333333333333           RACE/    INTACT: 1. WHI/AI/O 1. NO SCIEN
002 R/  12   (12         ) 122222122222122222           RACE/    INTACT: 1. WHI/AI/O 2. LIFESCI
003 R/  13   (13         ) 212222212222212222           RACE/    INTACT: 1. WHI/AI/O 3. PHYSSCI
004 R/  14   (14         ) 221222221222221222           RACE/    INTACT: 1. WHI/AI/O 4. EATHSCI
005 R/  15   (15         ) 222122222122222122           RACE/    INTACT: 1. WHI/AI/O 5. GEN SCI
006 R/  16   (16         ) 222212222212222212           RACE/    INTACT: 1. WHI/AI/O 6. INTESCI
007 R/  17   (17         ) 222221222221222221           RACE/    INTACT: 1. WHI/AI/O 7. MISSING
008 R/  21   (21         ) 111111222222222222           RACE/    INTACT: 2. BLACK    1. NO SCIEN
009 R/  22   (22         ) 322222222222222222           RACE/    INTACT: 2. BLACK    2. LIFESCI
010 R/  23   (23         ) 232222222222222222           RACE/    INTACT: 2. BLACK    3. PHYSSCI
011 R/  24   (24         ) 223222222222222222           RACE/    INTACT: 2. BLACK    4. EATHSCI
012 R/  25   (25         ) 222322222222222222           RACE/    INTACT: 2. BLACK    5. GEN SCI
013 R/  26   (26         ) 222232222222222222           RACE/    INTACT: 2. BLACK    6. INTESCI
014 R/  27   (27         ) 222223222222222222           RACE/    INTACT: 2. BLACK    7. MISSING
015 R/  31   (31         ) 222222111111222222           RACE/    INTACT: 3. HISPANIC 1. NO SCIEN
016 R/  32   (32         ) 222222322222222222           RACE/    INTACT: 3. HISPANIC 2. LIFESCI
017 R/  33   (33         ) 222222232222222222           RACE/    INTACT: 3. HISPANIC 3. PHYSSCI
018 R/  34   (34         ) 222222223222222222           RACE/    INTACT: 3. HISPANIC 4. EATHSCI
019 R/  35   (35         ) 222222222322222222           RACE/    INTACT: 3. HISPANIC 5. GEN SCI
020 R/  36   (36         ) 222222222232222222           RACE/    INTACT: 3. HISPANIC 6. INTESCI
021 R/  37   (37         ) 222222222223222222           RACE/    INTACT: 3. HISPANIC 7. MISSING
022 R/  41   (41         ) 222222222222111111           RACE/    INTACT: 4. ASIAN    1. NO SCIEN
023 R/  42   (42         ) 222222222222322222           RACE/    INTACT: 4. ASIAN    2. LIFESCI
024 R/  43   (43         ) 222222222222232222           RACE/    INTACT: 4. ASIAN    3. PHYSSCI
025 R/  44   (44         ) 222222222222223222           RACE/    INTACT: 4. ASIAN    4. EATHSCI
026 R/  45   (45         ) 222222222222222322           RACE/    INTACT: 4. ASIAN    5. GEN SCI
027 R/  46   (46         ) 222222222222222232           RACE/    INTACT: 4. ASIAN    6. INTESCI
028 R/  47   (47         ) 222222222222222223           RACE/    INTACT: 4. ASIAN    7. MISSING

CONDITIONING VARIABLE ID:  BACK0064
DESCRIPTION:               INTERACTION:  PARENTS' EDUCATION BY SCIENCE COURSES TAKING THIS YEAR
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:    PARE/
NAEP ID:                   N/A                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    35
TYPE OF CONTRAST:          INTERACTION                  NUMBER OF INDEPENDENT CONTRASTS:        24

001 P/  11   (11         ) 333333333333333333333333     PARE/    INTACT: 1. < HS     1. NO SCIEN
002 P/  12   (12         ) 122222122222122222122222     PARE/    INTACT: 1. < HS     2. LIFESCI
003 P/  13   (13         ) 212222212222212222212222     PARE/    INTACT: 1. < HS     3. PHYSSCI
004 P/  14   (14         ) 221222221222221222221222     PARE/    INTACT: 1. < HS     4. EATHSCI
005 P/  15   (15         ) 222122222122222122222122     PARE/    INTACT: 1. < HS     5. GEN SCI
006 P/  16   (16         ) 222212222212222212222212     PARE/    INTACT: 1. < HS     6. INTESCI
007 P/  17   (17         ) 222221222221222221222221     PARE/    INTACT: 1. < HS     7. MISSING
008 P/  21   (21         ) 111111222222222222222222     PARE/    INTACT: 2. HS GRAD  1. NO SCIEN
009 P/  22   (22         ) 322222222222222222222222     PARE/    INTACT: 2. HS GRAD  2. LIFESCI
010 P/  23   (23         ) 232222222222222222222222     PARE/    INTACT: 2. HS GRAD  3. PHYSSCI
011 P/  24   (24         ) 223222222222222222222222     PARE/    INTACT: 2. HS GRAD  4. EATHSCI
012 P/  25   (25         ) 222322222222222222222222     PARE/    INTACT: 2. HS GRAD  5. GEN SCI
013 P/  26   (26         ) 222232222222222222222222     PARE/    INTACT: 2. HS GRAD  6. INTESCI
014 P/  27   (27         ) 222223222222222222222222     PARE/    INTACT: 2. HS GRAD  7. MISSING
015 P/  31   (31         ) 222222111111222222222222     PARE/    INTACT: 3. POST HS  1. NO SCIEN
016 P/  32   (32         ) 222222322222222222222222     PARE/    INTACT: 3. POST HS  2. LIFESCI
017 P/  33   (33         ) 222222232222222222222222     PARE/    INTACT: 3. POST HS  3. PHYSSCI
018 P/  34   (34         ) 222222223222222222222222     PARE/    INTACT: 3. POST HS  4. EATHSCI
019 P/  35   (35         ) 222222222322222222222222     PARE/    INTACT: 3. POST HS  5. GEN SCI
020 P/  36   (36         ) 222222222232222222222222     PARE/    INTACT: 3. POST HS  6. INTESCI
021 P/  37   (37         ) 222222222223222222222222     PARE/    INTACT: 3. POST HS  7. MISSING
022 P/  41   (41         ) 222222222222111111222222     PARE/    INTACT: 4. COL GRAD 1. NO SCIEN
023 P/  42   (42         ) 222222222222322222222222     PARE/    INTACT: 4. COL GRAD 2. LIFESCI
024 P/  43   (43         ) 222222222222232222222222     PARE/    INTACT: 4. COL GRAD 3. PHYSSCI
025 P/  44   (44         ) 222222222222223222222222     PARE/    INTACT: 4. COL GRAD 4. EATHSCI
026 P/  45   (45         ) 222222222222222322222222     PARE/    INTACT: 4. COL GRAD 5. GEN SCI
027 P/  46   (46         ) 222222222222222232222222     PARE/    INTACT: 4. COL GRAD 6. INTESCI
028 P/  47   (47         ) 222222222222222223222222     PARE/    INTACT: 4. COL GRAD 7. MISSING
029 P/  51   (51         ) 222222222222222222111111     PARE/    INTACT: 5. PARED-?  1. NO SCIEN
030 P/  52   (52         ) 222222222222222222322222     PARE/    INTACT: 5. PARED-?  2. LIFESCI
031 P/  53   (53         ) 222222222222222222232222     PARE/    INTACT: 5. PARED-?  3. PHYSSCI
032 P/  54   (54         ) 222222222222222222223222     PARE/    INTACT: 5. PARED-?  4. EATHSCI
033 P/  55   (55         ) 222222222222222222222322     PARE/    INTACT: 5. PARED-?  5. GEN SCI
034 P/  56   (56         ) 222222222222222222222232     PARE/    INTACT: 5. PARED-?  6. INTESCI
035 P/  57   (57         ) 222222222222222222222223     PARE/    INTACT: 5. PARED-?  7. MISSING

CONDITIONING VARIABLE ID:  BACK0066
DESCRIPTION:               INTERACTION:  TYPE OF LOCALE (5 CATEGORIES) BY SCIENCE COURSES TAKING THIS YEAR
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:    TOL5/
NAEP ID:                   N/A                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    35
TYPE OF CONTRAST:          INTERACTION                  NUMBER OF INDEPENDENT CONTRASTS:        24

001 T/  11   (11         ) 333333333333333333333333     TOL5/    INTACT: 1. BIG CTY5 1. NO SCIEN
002 T/  12   (12         ) 122222122222122222122222     TOL5/    INTACT: 1. BIG CTY5 2. LIFESCI
003 T/  13   (13         ) 212222212222212222212222     TOL5/    INTACT: 1. BIG CTY5 3. PHYSSCI
004 T/  14   (14         ) 221222221222221222221222     TOL5/    INTACT: 1. BIG CTY5 4. EATHSCI
005 T/  15   (15         ) 222122222122222122222122     TOL5/    INTACT: 1. BIG CTY5 5. GEN SCI
006 T/  16   (16         ) 222212222212222212222212     TOL5/    INTACT: 1. BIG CTY5 6. INTESCI
007 T/  17   (17         ) 222221222221222221222221     TOL5/    INTACT: 1. BIG CTY5 7. MISSING
```

```
008 T/  21   (21          ) 111111222222222222222222         TOL5/    INTACT: 2. MID CTY5 1. NO SCIEN
009 T/  22   (22          ) 322222222222222222222222         TOL5/    INTACT: 2. MID CTY5 2. LIFESCI
010 T/  23   (23          ) 232222222222222222222222         TOL5/    INTACT: 2. MID CTY5 3. PHYSSCI
011 T/  24   (24          ) 223222222222222222222222         TOL5/    INTACT: 2. MID CTY5 4. EATHSCI
012 T/  25   (25          ) 222322222222222222222222         TOL5/    INTACT: 2. MID CTY5 5. GEN SCI
013 T/  26   (26          ) 222232222222222222222222         TOL5/    INTACT: 2. MID CTY5 6. INTESCI
014 T/  27   (27          ) 222223222222222222222222         TOL5/    INTACT: 2. MID CTY5 7. MISSING
015 T/  31   (31          ) 222222111111222222222222         TOL5/    INTACT: 3. FR/BTWN5 1. NO SCIEN
016 T/  32   (32          ) 222222322222222222222222         TOL5/    INTACT: 3. FR/BTWN5 2. LIFESCI
017 T/  33   (33          ) 222222232222222222222222         TOL5/    INTACT: 3. FR/BTWN5 3. PHYSSCI
018 T/  34   (34          ) 222222223222222222222222         TOL5/    INTACT: 3. FR/BTWN5 4. EATHSCI
019 T/  35   (35          ) 222222222322222222222222         TOL5/    INTACT: 3. FR/BTWN5 5. GEN SCI
020 T/  36   (36          ) 222222222232222222222222         TOL5/    INTACT: 3. FR/BTWN5 6. INTESCI
021 T/  37   (37          ) 222222222223222222222222         TOL5/    INTACT: 3. FR/BTWN5 7. MISSING
022 T/  41   (41          ) 222222222222111111222222         TOL5/    INTACT: 4. SML TWN5 1. NO SCIEN
023 T/  42   (42          ) 222222222222322222222222         TOL5/    INTACT: 4. SML TWN5 2. LIFESCI
024 T/  43   (43          ) 222222222222232222222222         TOL5/    INTACT: 4. SML TWN5 3. PHYSSCI
025 T/  44   (44          ) 222222222222223222222222         TOL5/    INTACT: 4. SML TWN5 4. EATHSCI
026 T/  45   (45          ) 222222222222222322222222         TOL5/    INTACT: 4. SML TWN5 5. GEN SCI
027 T/  46   (46          ) 222222222222222232222222         TOL5/    INTACT: 4. SML TWN5 6. INTESCI
028 T/  47   (47          ) 222222222222222223222222         TOL5/    INTACT: 4. SML TWN5 7. MISSING
029 T/  51   (51          ) 222222222222222222111111         TOL5/    INTACT: 5. RURAL5   1. NO SCIEN
030 T/  52   (52          ) 222222222222222222322222         TOL5/    INTACT: 5. RURAL5   2. LIFESCI
031 T/  53   (53          ) 222222222222222222232222         TOL5/    INTACT: 5. RURAL5   3. PHYSSCI
032 T/  54   (54          ) 222222222222222222223222         TOL5/    INTACT: 5. RURAL5   4. EATHSCI
033 T/  55   (55          ) 222222222222222222222322         TOL5/    INTACT: 5. RURAL5   5. GEN SCI
034 T/  56   (56          ) 222222222222222222222232         TOL5/    INTACT: 5. RURAL5   6. INTESCI
035 T/  57   (57          ) 222222222222222222222223         TOL5/    INTACT: 5. RURAL5   7. MISSING

CONDITIONING VARIABLE ID:  BACK0068
DESCRIPTION:               INTERACTION:  SCHOOL TYPE BY SCIENCE COURSES TAKING THIS YEAR
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:    SCHT/
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    21
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:        12

001 S/  11   (11          ) -1-1-1-1-1-1-1-1-1-1-1-1         SCHT/    INTACT: 1. PUBLIC   1. NO SCIEN
002 S/  12   (12          ) -10000000000-10000000000         SCHT/    INTACT: 1. PUBLIC   2. LIFESCI
003 S/  13   (13          ) 00-10000000000-100000000         SCHT/    INTACT: 1. PUBLIC   3. PHYSSCI
004 S/  14   (14          ) 0000-10000000000-1000000         SCHT/    INTACT: 1. PUBLIC   4. EATHSCI
005 S/  15   (15          ) 000000-10000000000-10000         SCHT/    INTACT: 1. PUBLIC   5. GEN SCI
006 S/  16   (16          ) 00000000-10000000000-100         SCHT/    INTACT: 1. PUBLIC   6. INTESCI
007 S/  17   (17          ) 0000000000-10000000000-1         SCHT/    INTACT: 1. PUBLIC   7. MISSING
008 S/  21   (21          ) -1-1-1-1-1-1000000000000         SCHT/    INTACT: 2. PRIVATE  1. NO SCIEN
009 S/  22   (22          ) 010000000000000000000000         SCHT/    INTACT: 2. PRIVATE  2. LIFESCI
010 S/  23   (23          ) 000100000000000000000000         SCHT/    INTACT: 2. PRIVATE  3. PHYSSCI
011 S/  24   (24          ) 000001000000000000000000         SCHT/    INTACT: 2. PRIVATE  4. EATHSCI
012 S/  25   (25          ) 000000010000000000000000         SCHT/    INTACT: 2. PRIVATE  5. GEN SCI
013 S/  26   (26          ) 000000000100000000000000         SCHT/    INTACT: 2. PRIVATE  6. INTESCI
014 S/  27   (27          ) 000000000001000000000000         SCHT/    INTACT: 2. PRIVATE  7. MISSING
015 S/  31   (31          ) 000000000000010101010101         SCHT/    INTACT: 3. CATHOLIC 1. NO SCIEN
016 S/  32   (32          ) 000000000000010000000000         SCHT/    INTACT: 3. CATHOLIC 2. LIFESCI
017 S/  33   (33          ) 000000000000000100000000         SCHT/    INTACT: 3. CATHOLIC 3. PHYSSCI
018 S/  34   (34          ) 000000000000000001000000         SCHT/    INTACT: 3. CATHOLIC 4. EATHSCI
019 S/  35   (35          ) 000000000000000000010000         SCHT/    INTACT: 3. CATHOLIC 5. GEN SCI
020 S/  36   (36          ) 000000000000000000000100         SCHT/    INTACT: 3. CATHOLIC 6. INTESCI
021 S/  37   (37          ) 000000000000000000000001         SCHT/    INTACT: 3. CATHOLIC 7. MISSING

CONDITIONING VARIABLE ID:  BACK0073
DESCRIPTION:               SAMPLE TYPE
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   SUBSAMP                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         2

001 SAMP S1  (01          ) 00                               SAMPLE S1
002 SAMP S2  (02          ) 10                               SAMPLE S2  AND S3
003 SAMP S3  (03          ) 01                               SAMPLE S3  AND S3

CONDITIONING VARIABLE ID:  BACK0074
DESCRIPTION:               INTERACTION:  SAMPLE BY RACE/ETHNICITY
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:        /RAC
NAEP ID:                   N/A                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    12
TYPE OF CONTRAST:          INTERACTION                       NUMBER OF INDEPENDENT CONTRASTS:         6

001  /R 11   (11          ) 010101010101                         /RAC INTACT: 1. SAMP S1  1. WHI/AI/O
002  /R 12   (12          ) -10000-10000                         /RAC INTACT: 1. SAMP S1  2. BLACK
003  /R 13   (13          ) 00-10000-100                         /RAC INTACT: 1. SAMP S1  3. HISPANIC
004  /R 14   (14          ) 0000-10000-1                         /RAC INTACT: 1. SAMP S1  4. ASIAN
005  /R 21   (21          ) -1-1-1000000                         /RAC INTACT: 2. SAMP S2  1. WHI/AI/O
006  /R 22   (22          ) 010000000000                         /RAC INTACT: 2. SAMP S2  2. BLACK
007  /R 23   (23          ) 000100000000                         /RAC INTACT: 2. SAMP S2  3. HISPANIC
008  /R 24   (24          ) 000001000000                         /RAC INTACT: 2. SAMP S2  4. ASIAN
009  /R 31   (31          ) 000000-1-1-1                         /RAC INTACT: 3. SAMP S3  1. WHI/AI/O
010  /R 32   (32          ) 000000010000                         /RAC INTACT: 3. SAMP S3  2. BLACK
011  /R 33   (33          ) 000000000100                         /RAC INTACT: 3. SAMP S3  3. HISPANIC
```

```
012  /R 34   (34          ) 000000000001                    /RAC INTACT: 3. SAMP S3  4. ASIAN

CONDITIONING VARIABLE ID:  BACK0075
DESCRIPTION:               DO YOU HAVE YOUR OWN STUDY DESK OR TABLE AT HOME?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   B008901                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        3

001 B008901Y (01         ) 000                             YES
002 B008901N (02         ) 100                             NO
003 B008901C (03         ) 010                             IDK
004 B008901M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  BACK0076
DESCRIPTION:               HOW SAFE DO YOU FEEL AT SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   B009401                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 VRY SAFE (01         ) 0000                            VERY SAFE
002 SAFE     (02         ) 1000                            SOMEWHAT SAFE
003 UNSAFE   (03         ) 0100                            SOMEWHAT UNSAFE
004 VRUNSAFE (04         ) 0010                            VERY UNSAFE
005 MISSING  (M          ) 0001                            MISSING

CONDITIONING VARIABLE ID:  BACK0077
DESCRIPTION:               DOES MOTHER OR STEPMOTHER LIVE AT HOME WITH YOU?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    MOM@HOME
NAEP ID:                   B005601                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        2

001 MOMHOM-Y (1          ) 00                              MOTHER AT HOME:  YES
002 MOMHOM-N (2          ) 10                              MOTHER AT HOME:  NO
003 MOMHOM-? (M          ) 01                              MOTHER AT HOME:  MISSING

CONDITIONING VARIABLE ID:  BACK0078
DESCRIPTION:               DOES FATHER OR STEPFATHER LIVE AT HOME WITH YOU?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:    DAD@HOME
NAEP ID:                   B005701                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        2

001 DADHOM-Y (1          ) 00                              FATHER AT HOME:  YES
002 DADHOM-N (2          ) 10                              FATHER AT HOME:  NO
003 DADHOM-? (M          ) 01                              FATHER AT HOME:  MISSING

CONDITIONING VARIABLE ID:  SUBJ0002
DESCRIPTION:               AGREE/DISAGREE: I LIKE SCIENCE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811001                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811001A (01         ) 000                             AGREE
002 K811001B (02         ) 100                             NOT SURE
003 K811001C (03         ) 010                             DISAGREE
004 K811001M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0003
DESCRIPTION:               AGREE/DISAGREE: I AM GOOD AT SCIENCE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811002                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811002A (01         ) 000                             AGREE
002 K811002B (02         ) 100                             NOT SURE
003 K811002C (03         ) 010                             DISAGREE
004 K811002M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0004
DESCRIPTION:               AGREE/DISAGREE: LEARNING SCI MOSTLY MEMORIZATION
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811003                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811003A (01         ) 000                             AGREE
002 K811003B (02         ) 100                             NOT SURE
003 K811003C (03         ) 010                             DISAGREE
004 K811003M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0005
DESCRIPTION:               AGREE/DISAGREE: SCI USEFUL FOR EVERYDAY PROBLEMS
```

```
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811004                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811004A (01        ) 000                               AGREE
002 K811004B (02        ) 100                               NOT SURE
003 K811004C (03        ) 010                               DISAGREE
004 K811004M (M         ) 001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0006
DESCRIPTION:               AGREE/DISAGREE: IF CHOICE, WOULD NOT STUDY SCIENCE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811005                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811005A (01        ) 000                               AGREE
002 K811005B (02        ) 100                               NOT SURE
003 K811005C (03        ) 010                               DISAGREE
004 K811005M (M         ) 001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0007
DESCRIPTION:               AGREE/DISAGREE: ALL CAN DO WELL IN SCI IF THEY TRY
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811006                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811006A (01        ) 000                               AGREE
002 K811006B (02        ) 100                               NOT SURE
003 K811006C (03        ) 010                               DISAGREE
004 K811006M (M         ) 001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0008
DESCRIPTION:               AGREE/DISAGREE: SCIENCE IS BORING
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811007                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811007A (01        ) 000                               AGREE
002 K811007B (02        ) 100                               NOT SURE
003 K811007C (03        ) 010                               DISAGREE
004 K811007M (M         ) 001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0009
DESCRIPTION:               AGREE/DISAGREE: SCIENCE IS A HARD SUBJECT
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811008                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811008A (01        ) 000                               AGREE
002 K811008B (02        ) 100                               NOT SURE
003 K811008C (03        ) 010                               DISAGREE
004 K811008M (M         ) 001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0010
DESCRIPTION:               EVER DONE HANDS-ON PROJECT WITH LIVING THINGS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811101                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811101Y (01        ) 0                                 YES
002 K811101M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SUBJ0011
DESCRIPTION:               EVER DONE HANDS-ON PROJECT WITH ELECTRICITY?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811102                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811102Y (01        ) 0                                 YES
002 K811102M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SUBJ0012
DESCRIPTION:               EVER DONE HANDS-ON PROJECT WITH CHEMICALS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811103                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811103Y (01        ) 0                                 YES
002 K811103M (M         ) 1                                 MISSING
```

```
CONDITIONING VARIABLE ID:  SUBJ0013
DESCRIPTION:               EVER DONE HANDS-ON PROJECT WITH ROCKS OR MINERALS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811104                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811104Y (01         ) 0                               YES
002 K811104M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0014
DESCRIPTION:               DONE HANDS-ON PROJ W/ MAGNIFYING GLASS/MICROSCOPE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811105                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811105Y (01         ) 0                               YES
002 K811105M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0015
DESCRIPTION:               DONE HANDS-ON PROJ W/ THERMOMETER OR BAROMETER?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811106                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811106Y (01         ) 0                               YES
002 K811106M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0016
DESCRIPTION:               EVER DONE HANDS-ON PROJECT WITH SIMPLE MACHINES?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811107                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811107Y (01         ) 0                               YES
002 K811107M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0017
DESCRIPTION:               HAVE DONE HANDS-ON PROJECT WITH NONE OF THE ABOVE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811108                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 K811108Y (01         ) 0                               YES
002 K811108M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0018
DESCRIPTION:               HOW OFTEN DO YOU STUDY SCIENCE IN SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811201                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        5

001 K811201A (01         ) 00000                           EVERY DAY
002 K811201B (02         ) 10000                           3-4 TIMES A WEEK
003 K811201C (03         ) 01000                           1-2 TIMES A WEEK
004 K811201D (04         ) 00100                           LESS THEN ONCE WEEK
005 K811201E (05         ) 00010                           NEVER
006 K811201M (M          ) 00001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0020
DESCRIPTION:               DO SCI PROJECTS IN SCHOOL THAT TAKE 1 OR MORE WKS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K811401                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        2

001 K811401Y (01         ) 00                              YES
002 K811401N (02         ) 10                              NO
003 K811401M (M          ) 01                              MISSING

CONDITIONING VARIABLE ID:  SUBJ0021
DESCRIPTION:               LAST 2 YRS, BEEN IN SCI FAIR, FESTIVAL, SCI DAY?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811501                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        2

001 K811501Y (01         ) 00                              YES
002 K811501N (02         ) 10                              NO
003 K811501M (M          ) 01                              MISSING
```

```
CONDITIONING VARIABLE ID:  SUBJ0022
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN DO YOU READ TEXTBOOK?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811601                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811601A (01        ) 0000                             ALMOST EVERY DAY
002 K811601B (02        ) 1000                             ONCE OR TWICE A WEEK
003 K811601C (03        ) 0100                             ONCE OR TWICE A MTH
004 K811601D (04        ) 0010                             NEVER OR HARDLY EVER
005 K811601M (M         ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0023
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN DO YOU READ MAGS/BKS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811602                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811602A (01        ) 0000                             ALMOST EVERY DAY
002 K811602B (02        ) 1000                             ONCE OR TWICE A WEEK
003 K811602C (03        ) 0100                             ONCE OR TWICE A MTH
004 K811602D (04        ) 0010                             NEVER OR HARDLY EVER
005 K811602M (M         ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0024
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN DISCUSS SCIENCE NEWS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811603                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811603A (01        ) 0000                             ALMOST EVERY DAY
002 K811603B (02        ) 1000                             ONCE OR TWICE A WEEK
003 K811603C (03        ) 0100                             ONCE OR TWICE A MTH
004 K811603D (04        ) 0010                             NEVER OR HARDLY EVER
005 K811603M (M         ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0025
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN WORK WITH OTHERS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811604                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811604A (01        ) 0000                             ALMOST EVERY DAY
002 K811604B (02        ) 1000                             ONCE OR TWICE A WEEK
003 K811604C (03        ) 0100                             ONCE OR TWICE A MTH
004 K811604D (04        ) 0010                             NEVER OR HARDLY EVER
005 K811604M (M         ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0026
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN GIVE ORAL REPORT?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811605                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811605A (01        ) 0000                             ALMOST EVERY DAY
002 K811605B (02        ) 1000                             ONCE OR TWICE A WEEK
003 K811605C (03        ) 0100                             ONCE OR TWICE A MTH
004 K811605D (04        ) 0010                             NEVER OR HARDLY EVER
005 K811605M (M         ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0027
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN GIVE WRITTEN REPORT?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811606                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811606A (01        ) 0000                             ALMOST EVERY DAY
002 K811606B (02        ) 1000                             ONCE OR TWICE A WEEK
003 K811606C (03        ) 0100                             ONCE OR TWICE A MTH
004 K811606D (04        ) 0010                             NEVER OR HARDLY EVER
005 K811606M (M         ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SUBJ0030
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN DO YOU USE COMPUTER?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811609                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811609A (01        ) 0000                             ALMOST EVERY DAY
002 K811609B (02        ) 1000                             ONCE OR TWICE A WEEK
```

```
003 K811609C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811609D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811609M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0031
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN TAKE TEST OR QUIZ?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811610                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811610A (01         ) 0000                           ALMOST EVERY DAY
002 K811610B (02         ) 1000                           ONCE OR TWICE A WEEK
003 K811610C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811610D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811610M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0032
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN DO YOU USE LIBRARY?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811611                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811611A (01         ) 0000                           ALMOST EVERY DAY
002 K811611B (02         ) 1000                           ONCE OR TWICE A WEEK
003 K811611C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811611D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811611M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0033
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN OBSERVE/MEAS OUTSIDE?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811612                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811612A (01         ) 0000                           ALMOST EVERY DAY
002 K811612B (02         ) 1000                           ONCE OR TWICE A WEEK
003 K811612C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811612D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811612M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0034
DESCRIPTION:               HOW OFTEN DOES SCIENCE TEACHER TALK TO CLASS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811701                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811701A (01         ) 0000                           ALMOST EVERY DAY
002 K811701B (02         ) 1000                           ONCE OR TWICE A WEEK
003 K811701C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811701D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811701M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0035
DESCRIPTION:               HOW OFTEN DOES SCIENCE TEACHER DO DEMONSTRATION?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811702                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811702A (01         ) 0000                           ALMOST EVERY DAY
002 K811702B (02         ) 1000                           ONCE OR TWICE A WEEK
003 K811702C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811702D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811702M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0036
DESCRIPTION:               HOW OFTEN DOES SCIENCE TEACHER SHOW VIDEO OR TV?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811703                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811703A (01         ) 0000                           ALMOST EVERY DAY
002 K811703B (02         ) 1000                           ONCE OR TWICE A WEEK
003 K811703C (03         ) 0100                           ONCE OR TWICE A MTH
004 K811703D (04         ) 0010                           NEVER OR HARDLY EVER
005 K811703M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SUBJ0037
DESCRIPTION:               HOW OFTEN DOES SCIENCE TEACHER USE COMPUTER?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811704                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
```

```
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811704A (01         ) 0000                               ALMOST EVERY DAY
002 K811704B (02         ) 1000                               ONCE OR TWICE A WEEK
003 K811704C (03         ) 0100                               ONCE OR TWICE A MTH
004 K811704D (04         ) 0010                               NEVER OR HARDLY EVER
005 K811704M (M          ) 0001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0038
DESCRIPTION:               HOW OFTEN DOES SCI TEACHER USE CD'S/LASER DISCS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811705                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        4

001 K811705A (01         ) 0000                               ALMOST EVERY DAY
002 K811705B (02         ) 1000                               ONCE OR TWICE A WEEK
003 K811705C (03         ) 0100                               ONCE OR TWICE A MTH
004 K811705D (04         ) 0010                               NEVER OR HARDLY EVER
005 K811705M (M          ) 0001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0039
DESCRIPTION:               HOW OFTEN DOES SCI CLASS GO ON A FIELD TRIP?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811801                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811801A (01         ) 000                                3 OR MORE A DAY
002 K811801B (02         ) 100                                1 OR 2 TIMES A YEAR
003 K811801C (03         ) 010                                NEVER OR HARDLY EVER
004 K811801M (M          ) 001                                MISSING

CONDITIONING VARIABLE ID:  SUBJ0040
DESCRIPTION:               HOW OFTEN DOES GUEST SPEAKER COME TO SCI CLASS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811901                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        3

001 K811901A (01         ) 000                                3 OR MORE A DAY
002 K811901B (02         ) 100                                1 OR 2 TIMES A YEAR
003 K811901C (03         ) 010                                NEVER OR HARDLY EVER
004 K811901M (M          ) 001                                MISSING

CONDITIONING VARIABLE ID:  SUBJ0041
DESCRIPTION:               ABOUT HOW MANY QUESTIONS RIGHT ON TEST?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   SM00101                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        4

001 SM00101A (01         ) 0000                               ALMOST ALL
002 SM00101B (02         ) 1000                               MORE THAN HALF
003 SM00101C (03         ) 0100                               ABOUT HALF
004 SM00101D (04         ) 0010                               LESS THAN HALF
005 SM00101M (M          ) 0001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0042
DESCRIPTION:               HOW HARD TEST COMPARED TO THOSE IN SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   SM00201                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        4

001 SM00201A (01         ) 0000                               MUCH HARDER
002 SM00201B (02         ) 1000                               HARDER THAN OTHERS
003 SM00201C (03         ) 0100                               ABOUT THE SAME
004 SM00201D (04         ) 0010                               EASIER THEN OTHES
005 SM00201M (M          ) 0001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0043
DESCRIPTION:               HOW HARD DID YOU TRY ON TEST COMPARED TO OTHERS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   SM00301                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        4

001 SM00301A (01         ) 0000                               MUCH HARDER
002 SM00301B (02         ) 1000                               HARDER THAN OTHERS
003 SM00301C (03         ) 0100                               ABOUT AS HARD
004 SM00301D (04         ) 0010                               NOT AS HARD AS OTHER
005 SM00301M (M          ) 0001                               MISSING

CONDITIONING VARIABLE ID:  SUBJ0044
DESCRIPTION:               HOW IMPORTANT WAS IT YOU DO WELL ON THIS TEST?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
```

```
CONDITIONING VAR LABEL:
NAEP ID:                   SM00401                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 SM00401A (01         ) 0000                              VERY IMPORTANT
002 SM00401B (02         ) 1000                              IMPORTANT
003 SM00401C (03         ) 0100                              SOMEWHAT IMPORTANT
004 SM00401D (04         ) 0010                              NOT VERY IMPORTANTER
005 SM00401M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SUBJ0045
DESCRIPTION:               HOW OFTEN HAD TO WRITE LONG ANSWERS TO QSTS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   SM00501                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 SM00501A (01         ) 0000                              AT LEAST ONCE A WEEK
002 SM00501B (02         ) 1000                              ONCE OR TWICE A MNTH
003 SM00501C (03         ) 0100                              ONCE OR TWICE A YEAR
004 SM00501D (04         ) 0010                              NEVERERY IMPORTANTER
005 SM00501M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  BACK0079
DESCRIPTION:               DESCRIBE YOUR OVERALL GRADES SINCE 6TH GRADE
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   B009701                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    7
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        6

001 B009701A (01         ) 000000                            MOSTLY A'S
002 B009701B (02         ) 100000                            MOSTLY B'S
003 B009701C (03         ) 010000                            MOSTLY C'S
004 B009701D (04         ) 001000                            MOSTLY D'S
005 B009701E (05         ) 000100                            MOSTLY BELOW D'S
006 B009701F (06         ) 000010                            CLASSES NOT GRADED
007 B009701M (M          ) 000001                            MISSING

CONDITIONING VARIABLE ID:  BACK0080
DESCRIPTION:               HOW FAR IN SCHOOL DO YOU THINK YOU WILL GO?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   B009801                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    7
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        6

001 B009801A (01         ) 000000                            NOT FINISH HS
002 B009801B (02         ) 100000                            GRADUATE HS
003 B009801C (03         ) 010000                            SOME ED PAST HS
004 B009801D (04         ) 001000                            GRADUATE COLLEGE
005 B009801E (05         ) 000100                            GO GRADUATE SCHOOL
006 B009801F (06         ) 000010                            I DON'T KNOW
007 B009801M (M          ) 000001                            MISSING

CONDITIONING VARIABLE ID:  BACK0083
DESCRIPTION:               DOES YOUR STEP/MOTHER WORK AT A JOB FOR PAY?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   B009601                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 B009601A (01         ) 0000                              YES, FULL-TIME
002 B009601B (02         ) 1000                              YES, PART-TIME
003 B009601N (03         ) 0100                              NO
004 B009601D (04         ) 0010                              DON'T LIVE W/ MOTHER
005 B009601E (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  BACK0084
DESCRIPTION:               DOES YOUR STEP/FATHER WORK AT A JOB FOR PAY?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   B009602                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 B009602A (01         ) 0000                              YES, FULL-TIME
002 B009602B (02         ) 1000                              YES, PART-TIME
003 B009602N (03         ) 0100                              NO
004 B009602D (04         ) 0010                              DON'T LIVE W/ FATHER
005 B009602M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SUBJ0046
DESCRIPTION:               DO YOU/TEACHER SAVE YOUR SCI WORK IN A PORTFOLIO?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   K812101                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        2

001 K812101Y (01         ) 00                                YES
```

```
002 K812101N (02         ) 10                              NO
003 K812101M (M          ) 01                              MISSING

CONDITIONING VARIABLE ID:  SUBJ0047
DESCRIPTION:               HOW MUCH TIME WEEKLY SPENT ON SCIENCE HOMEWORK?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K812201                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     8
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         7

001 K812201A (01         ) 0000000                         NOT TAKING SCIENCE
002 K812201B (02         ) 1000000                         NONE
003 K812201C (03         ) 0100000                         1/2 HOUR
004 K812201D (04         ) 0010000                         1 HOUR
005 K812201E (05         ) 0001000                         2 HOURS
006 K812201F (06         ) 0000100                         3 HOURS
007 K812201G (07         ) 0000010                         MORE THAN 3 HOURS
008 K812201M (M          ) 0000001                         MISSING

CONDITIONING VARIABLE ID:  SUBJ0048
DESCRIPTION:               FOR SCI IN SCHOOL, HOW OFTEN HANDS-ON ACTIVITIES?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811613                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         4

001 K811613A (01         ) 0000                            ALMOST EVERY DAY
002 K811613B (02         ) 1000                            ONCE OR TWICE A WEEK
003 K811613C (03         ) 0100                            ONCE OR TWICE A MTH
004 K811613D (04         ) 0010                            NEVER OR HARDLY EVER
005 K811613M (M          ) 0001                            MISSING

CONDITIONING VARIABLE ID:  SUBJ0049
DESCRIPTION:               FOR SCI, HOW OFTEN DISCUSS HANDS-ON RESULTS?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811614                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         4

001 K811614A (01         ) 0000                            ALMOST EVERY DAY
002 K811614B (02         ) 1000                            ONCE OR TWICE A WEEK
003 K811614C (03         ) 0100                            ONCE OR TWICE A MTH
004 K811614D (04         ) 0010                            NEVER OR HARDLY EVER
005 K811614M (M          ) 0001                            MISSING

CONDITIONING VARIABLE ID:  SUBJ0050
DESCRIPTION:               FOR SCI, DESIGN & CARRY OUT OWN INVESTIGATION?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   K811615                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         4

001 K811615A (01         ) 0000                            ALMOST EVERY DAY
002 K811615B (02         ) 1000                            ONCE OR TWICE A WEEK
003 K811615C (03         ) 0100                            ONCE OR TWICE A MTH
004 K811615D (04         ) 0010                            NEVER OR HARDLY EVER
005 K811615M (M          ) 0001                            MISSING

CONDITIONING VARIABLE ID:  SCHL0009
DESCRIPTION:               HAS MATH BEEN IDENTIFIED AS A PRIORITY?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C031603                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         2

001 C031603Y (01         ) 00                              YES
002 C031603N (02         ) 10                              NO
003 C031603M (M          ) 01                              MISSING

CONDITIONING VARIABLE ID:  SCHL0010
DESCRIPTION:               HAS SCIENCE BEEN IDENTIFIED AS A PRIORITY?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C031607                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         2

001 C031607Y (01         ) 00                              YES
002 C031607N (02         ) 10                              NO
003 C031607M (M          ) 01                              MISSING

CONDITIONING VARIABLE ID:  SCHL0012
DESCRIPTION:               HAS ARTS BEEN IDENTIFIED AS A PRIORITY?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C031610                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         2
```

```
001 C031610Y (01         ) 00                             YES
002 C031610N (02         ) 10                             NO
003 C031610M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0013
DESCRIPTION:               HAS SUBJECT INTEGRATION BEEN A PRIORITY?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C031606                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        2

001 C031606Y (01         ) 00                             YES
002 C031606N (02         ) 10                             NO
003 C031606M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0014
DESCRIPTION:               COMPUTERS AVAILABLE ALL THE TIME IN CLASSROOM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C035701                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        2

001 C035701Y (01         ) 00                             YES
002 C035701N (02         ) 10                             NO
003 C035701M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0015
DESCRIPTION:               COMPUTERS GROUPED IN SEPARATE LAB AND AVAILABLE?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C035702                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        2

001 C035702Y (01         ) 00                             YES
002 C035702N (02         ) 10                             NO
003 C035702M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0016
DESCRIPTION:               COMPUTERS AVAILABLE TO BRING TO ROOM WHEN NEEDED?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C035703                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        2

001 C035703Y (01         ) 00                             YES
002 C035703N (02         ) 10                             NO
003 C035703M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0017
DESCRIPTION:               SCHOOL W/ SPECIAL FOCUS ON MATH?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037201                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037201Y (01        ) 0                               YES
002 C037201M (M         ) 1                               MISSING

CONDITIONING VARIABLE ID:  SCHL0018
DESCRIPTION:               SCHOOL W/ SPECIAL FOCUS ON SCIENCE?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037202                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037202Y (01        ) 0                               YES
002 C037202M (M         ) 1                               MISSING

CONDITIONING VARIABLE ID:  SCHL0020
DESCRIPTION:               SCHOOL W/ SPECIAL FOCUS ON ARTS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037204                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037204Y (01        ) 0                               YES
002 C037204M (M         ) 1                               MISSING

CONDITIONING VARIABLE ID:  SCHL0021
DESCRIPTION:               SCHOOL W/ SPECIAL FOCUS ON OTHER?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037205                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037205Y (01        ) 0                               YES
002 C037205M (M         ) 1                               MISSING
```

```
CONDITIONING VARIABLE ID:  SCHL0022
DESCRIPTION:               SCHOOL NOT A SPECIAL FOCUS SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037206                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037206Y (01         ) 0                              YES
002 C037206M (M          ) 1                              MISSING

CONDITIONING VARIABLE ID:  SCHL0023
DESCRIPTION:               SCHOOL FOLLOW DISTRICT/STATE MATH CURRICULUM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037301                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037301Y (01         ) 0                              YES
002 C037301M (M          ) 1                              MISSING

CONDITIONING VARIABLE ID:  SCHL0024
DESCRIPTION:               SCHOOL FOLLOW DISTRICT/STATE SCIENCE CURRICULUM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037302                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037302Y (01         ) 0                              YES
002 C037302M (M          ) 1                              MISSING

CONDITIONING VARIABLE ID:  SCHL0026
DESCRIPTION:               SCHOOL FOLLOW DISTRICT/STATE ARTS CURRICULUM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037304                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037304Y (01         ) 0                              YES
002 C037304M (M          ) 1                              MISSING

CONDITIONING VARIABLE ID:  SCHL0027
DESCRIPTION:               SCHOOL FOLLOW DISTRICT/STATE FOR NONE OF ABOVE?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037305                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037305Y (01         ) 0                              YES
002 C037305M (M          ) 1                              MISSING

CONDITIONING VARIABLE ID:  SCHL0044
DESCRIPTION:               WHICH BEST DESCRIBES PRIMARY WAY LIBRARY STAFFED?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C036601                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        4

001 C036601A (01         ) 0000                           NO LIBRARY IN SCHOOL
002 C036601B (02         ) 1000                           LIBRARY-NO/VOL STAFF
003 C036601C (03         ) 0100                           PART-TIME STAFF
004 C036601D (04         ) 0010                           FULL-TIME STAFF
005 C036601M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  SCHL0045
DESCRIPTION:               INVOLVE PARENTS AS AIDES IN CLASSROOM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032207                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        3

001 C032207A (01         ) 000                            YES, ROUTINELY
002 C032207B (02         ) 100                            YES, OCCASIONALLY
003 C032207N (03         ) 010                            NO
004 C032207M (M          ) 001                            MISSING

CONDITIONING VARIABLE ID:  SCHL0046
DESCRIPTION:               HAVE PARENTS REVIEW/SIGN HOMEWORK?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032209                        TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:        3

001 C032209A (01         ) 000                            YES, ROUTINELY
002 C032209B (02         ) 100                            YES, OCCASIONALLY
003 C032209N (03         ) 010                            NO
004 C032209M (M          ) 001                            MISSING
```

```
CONDITIONING VARIABLE ID:  SCHL0047
DESCRIPTION:               ASSIGN HOMEWORK STUDENTS DO WITH PARENTS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032210                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 C032210A (01         ) 000                              YES, ROUTINELY
002 C032210B (02         ) 100                              YES, OCCASIONALLY
003 C032210N (03         ) 010                              NO
004 C032210M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0048
DESCRIPTION:               HAVE A PARENT VOLUNTEER PROGRAM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032211                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 C032211A (01         ) 000                              YES, ROUTINELY
002 C032211B (02         ) 100                              YES, OCCASIONALLY
003 C032211N (03         ) 010                              NO
004 C032211M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0049
DESCRIPTION:               WHAT % OF PARENTS IN PARENT-TEACHER ORGS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037701                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C037701A (01         ) 0000                             0-25%
002 C037701B (02         ) 1000                             26-50%
003 C037701C (03         ) 0100                             51-75%
004 C037701D (04         ) 0010                             76-100%
005 C037701M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0050
DESCRIPTION:               WHAT % OF PARENTS IN OPEN HOUSE/BACK SCHOOL NIGHT?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037702                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C037702A (01         ) 0000                             0-25%
002 C037702B (02         ) 1000                             26-50%
003 C037702C (03         ) 0100                             51-75%
004 C037702D (04         ) 0010                             76-100%
005 C037702M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0051
DESCRIPTION:               WHAT % OF PARENTS IN PARENT-TEACHER COMFERENCES?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037703                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C037703A (01         ) 0000                             0-25%
002 C037703B (02         ) 1000                             26-50%
003 C037703C (03         ) 0100                             51-75%
004 C037703D (04         ) 0010                             76-100%
005 C037703M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0052
DESCRIPTION:               WHAT % PARENTS INVOLVED MAKING CURRICULUM DECISION
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037704                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C037704A (01         ) 0000                             0-25%
002 C037704B (02         ) 1000                             26-50%
003 C037704C (03         ) 0100                             51-75%
004 C037704D (04         ) 0010                             76-100%
005 C037704M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0053
DESCRIPTION:               WHAT % OF PARENTS IN VOLUNTEER PROGRAMS?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037705                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C037705A (01         ) 0000                             0-25%
002 C037705B (02         ) 1000                             26-50%
003 C037705C (03         ) 0100                             51-75%
```

```
004 C037705D (04         ) 0010                              76-100%
005 C037705M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0054
DESCRIPTION:               IS STUDENT ABSENTEEISM A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032402                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032402A (01         ) 0000                              SERIOUS
002 C032402B (02         ) 1000                              MODERATE
003 C032402C (03         ) 0100                              MINOR
004 C032402D (04         ) 0010                              NOT A PROBLEM
005 C032402M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0055
DESCRIPTION:               IS STUDENT TARDINESS A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032401                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032401A (01         ) 0000                              SERIOUS
002 C032401B (02         ) 1000                              MODERATE
003 C032401C (03         ) 0100                              MINOR
004 C032401D (04         ) 0010                              NOT A PROBLEM
005 C032401M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0056
DESCRIPTION:               ARE PHYSICAL CONFLICTS A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032404                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032404A (01         ) 0000                              SERIOUS
002 C032404B (02         ) 1000                              MODERATE
003 C032404C (03         ) 0100                              MINOR
004 C032404D (04         ) 0010                              NOT A PROBLEM
005 C032404M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0057
DESCRIPTION:               IS TEACHER ABSENTEEISM A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032406                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032406A (01         ) 0000                              SERIOUS
002 C032406B (02         ) 1000                              MODERATE
003 C032406C (03         ) 0100                              MINOR
004 C032406D (04         ) 0010                              NOT A PROBLEM
005 C032406M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0058
DESCRIPTION:               ARE RACE/CULT. CONFLICTS A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032407                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032407A (01         ) 0000                              SERIOUS
002 C032407B (02         ) 1000                              MODERATE
003 C032407C (03         ) 0100                              MINOR
004 C032407D (04         ) 0010                              NOT A PROBLEM
005 C032407M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0059
DESCRIPTION:               IS STUDENT HEALTH A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032408                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032408A (01         ) 0000                              SERIOUS
002 C032408B (02         ) 1000                              MODERATE
003 C032408C (03         ) 0100                              MINOR
004 C032408D (04         ) 0010                              NOT A PROBLEM
005 C032408M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0060
DESCRIPTION:               IS LACK OF PARENT INVLMT A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032409                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4
```

```
001 C032409A (01         ) 0000                          SERIOUS
002 C032409B (02         ) 1000                          MODERATE
003 C032409C (03         ) 0100                          MINOR
004 C032409D (04         ) 0010                          NOT A PROBLEM
005 C032409M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0061
DESCRIPTION:               IS STUD USE OF ALCOHOL A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032410                       TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                         NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032410A (01         ) 0000                          SERIOUS
002 C032410B (02         ) 1000                          MODERATE
003 C032410C (03         ) 0100                          MINOR
004 C032410D (04         ) 0010                          NOT A PROBLEM
005 C032410M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0062
DESCRIPTION:               IS STUDENT TOBACCO USE A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032411                       TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                         NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032411A (01         ) 0000                          SERIOUS
002 C032411B (02         ) 1000                          MODERATE
003 C032411C (03         ) 0100                          MINOR
004 C032411D (04         ) 0010                          NOT A PROBLEM
005 C032411M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0063
DESCRIPTION:               IS STUDENT DRUG USE A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032412                       TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                         NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032412A (01         ) 0000                          SERIOUS
002 C032412B (02         ) 1000                          MODERATE
003 C032412C (03         ) 0100                          MINOR
004 C032412D (04         ) 0010                          NOT A PROBLEM
005 C032412M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0064
DESCRIPTION:               ARE GANG ACTIVITIES A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032413                       TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                         NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032413A (01         ) 0000                          SERIOUS
002 C032413B (02         ) 1000                          MODERATE
003 C032413C (03         ) 0100                          MINOR
004 C032413D (04         ) 0010                          NOT A PROBLEM
005 C032413M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0065
DESCRIPTION:               IS STUDENT MISBEHAVIOR A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032414                       TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                         NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032414A (01         ) 0000                          SERIOUS
002 C032414B (02         ) 1000                          MODERATE
003 C032414C (03         ) 0100                          MINOR
004 C032414D (04         ) 0010                          NOT A PROBLEM
005 C032414M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0066
DESCRIPTION:               IS STUDENT CHEATING A PROBLEM IN YOUR SCHOOL?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032415                       TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                         NUMBER OF INDEPENDENT CONTRASTS:         4

001 C032415A (01         ) 0000                          SERIOUS
002 C032415B (02         ) 1000                          MODERATE
003 C032415C (03         ) 0100                          MINOR
004 C032415D (04         ) 0010                          NOT A PROBLEM
005 C032415M (M          ) 0001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0067
DESCRIPTION:               IS TEACHER MORALE POS. OR NEG.?
```

```
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032502                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C032502A (01         ) 0000                             VERY POSITIVE
002 C032502B (02         ) 1000                             SOMEWHAT POSITIVE
003 C032502C (03         ) 0100                             SOMEWHAT NEGATIVE
004 C032502D (04         ) 0010                             VERY NEGATIVE
005 C032502M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0068
DESCRIPTION:               ARE STUDENT ATTITUDES TO ACADEMICS POS. OR NEG.?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032503                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C032503A (01         ) 0000                             VERY POSITIVE
002 C032503B (02         ) 1000                             SOMEWHAT POSITIVE
003 C032503C (03         ) 0100                             SOMEWHAT NEGATIVE
004 C032503D (04         ) 0010                             VERY NEGATIVE
005 C032503M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0069
DESCRIPTION:               IS PARENT SUPPORT FOR ACHIEVEMENT POS. OR NEG.?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032505                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C032505A (01         ) 0000                             VERY POSITIVE
002 C032505B (02         ) 1000                             SOMEWHAT POSITIVE
003 C032505C (03         ) 0100                             SOMEWHAT NEGATIVE
004 C032505D (04         ) 0010                             VERY NEGATIVE
005 C032505M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0070
DESCRIPTION:               IS REGARD FOR SCHOOL PROPERTY POS. OR NEG.?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C032506                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C032506A (01         ) 0000                             VERY POSITIVE
002 C032506B (02         ) 1000                             SOMEWHAT POSITIVE
003 C032506C (03         ) 0100                             SOMEWHAT NEGATIVE
004 C032506D (04         ) 0010                             VERY NEGATIVE
005 C032506M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0071
DESCRIPTION:               % ABSENT ON AVERAGE DAY?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C033601                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C033601A (01         ) 0000                             0-2%
002 C033601B (02         ) 1000                             3-5%
003 C033601C (03         ) 0100                             6-10%
004 C033601D (04         ) 0010                             MORE THAN 10%
005 C033601M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0072
DESCRIPTION:               WHAT % OF TEACHERS ABSENT ON GIVEN DAY?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C036501                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 C036501A (01         ) 0000                             0-2%
002 C036501B (02         ) 1000                             3-5%
003 C036501C (03         ) 0100                             6-10%
004 C036501D (04         ) 0010                             MORE THAN 10%
005 C036501M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  SCHL0073
DESCRIPTION:               % OF STUDS EROLLED AT START OF YR EROLLED AT END?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037801                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    9
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        8

001 C037801A (01         ) 00000000                         98-100%
002 C037801B (02         ) 10000000                         95-97%
003 C037801C (03         ) 01000000                         90-94%
004 C037801D (04         ) 00100000                         80-89%
```

```
005 C037801E (05         ) 00010000                          70-79%
006 C037801F (06         ) 00001000                          60-69%
007 C037801G (07         ) 00000100                          50-59%
008 C037801H (08         ) 00000010                          LESS THAN 50%
009 C037801M (M          ) 00000001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0075
DESCRIPTION:               % OF FULL TIME TEACHERS LEFT BEFORE END OF YR?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C038001                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     8
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         7

001 C038001A (01         ) 0000000                           0%
002 C038001B (02         ) 1000000                           1-2%
003 C038001C (03         ) 0100000                           3-5%
004 C038001D (04         ) 0010000                           6-10%
005 C038001E (05         ) 0001000                           11-15%
006 C038001F (06         ) 0000100                           16-20%
007 C038001G (07         ) 0000010                           MORE THAN 20%
008 C038001M (M          ) 0000001                           MISSING

CONDITIONING VARIABLE ID:  SCHL0076
DESCRIPTION:               IS SCHOOL IN NATIONAL SCHOOL LUNCH PROGRAM?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C038301                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         2

001 C038301Y (01         ) 00                                YES
002 C038301N (02         ) 10                                NO
003 C038301M (M          ) 01                                MISSING

CONDITIONING VARIABLE ID:  SCHL0077
DESCRIPTION:               SCHOOL RECEIVE CHAP 1/TITLE 1 FUNDING?
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C038801                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         2

001 C038801Y (01         ) 00                                YES
002 C038801N (02         ) 10                                NO
003 C038801M (M          ) 01                                MISSING

CONDITIONING VARIABLE ID:  SCHL0078
DESCRIPTION:               DID PRINCIPAL FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034101                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         1

001 C034101Y (01         ) 0                                 YES
002 C034101M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0079
DESCRIPTION:               DID HEADMASTER/HEADMISTRESS FILL OUT QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034102                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         1

001 C034102Y (01         ) 0                                 YES
002 C034102M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0080
DESCRIPTION:               DID HEAD TEACHER FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034103                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         1

001 C034103Y (01         ) 0                                 YES
002 C034103M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0081
DESCRIPTION:               DID VICE PRINCIPAL FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034104                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         1

001 C034104Y (01         ) 0                                 YES
002 C034104M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0082
DESCRIPTION:               DID COUNSELOR FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
```

```
CONDITIONING VAR LABEL:
NAEP ID:                   C034105                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        1

001 C034105Y (01         ) 0                                 YES
002 C034105M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0083
DESCRIPTION:               DID CURRICULUM COORD FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034106                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        1

001 C034106Y (01         ) 0                                 YES
002 C034106M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0084
DESCRIPTION:               DID TEACHER FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034107                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        1

001 C034107Y (01         ) 0                                 YES
002 C034107M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0085
DESCRIPTION:               DID SECRETARY FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034108                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        1

001 C034108Y (01         ) 0                                 YES
002 C034108M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  SCHL0086
DESCRIPTION:               DID OTHER PERSON FILL OUT THIS QUESTIONNAIRE
GRADES/ASSESSMENTS:        N04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C034109                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        1

001 C034109Y (01         ) 0                                 YES
002 C034109M (M          ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0001
DESCRIPTION:               WHAT IS YOUR GENDER?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T055901                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        2

001 T055901A (01         ) 00                                MALE
002 T055901B (02         ) 10                                FEMALE
003 T055901M (M          ) 01                                MISSING

CONDITIONING VARIABLE ID:  TCHR0002
DESCRIPTION:               WHICH BEST DESRIBES YOU?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056001                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    7
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        6

001 T056001A (01         ) 000000                            WHITE
002 T056001B (02         ) 100000                            BLACK
003 T056001C (03         ) 010000                            HISPANIC
004 T056001D (04         ) 001000                            ASIAN/PACIFIC AMERIC
005 T056001E (05         ) 000100                            AMER IND/ALASKA NATV
006 T056001F (06         ) 000010                            OTHER
007 T056001M (M          ) 000001                            MISSING

CONDITIONING VARIABLE ID:  TCHR0003
DESCRIPTION:               YEARS TAUGHT
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040301                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        5

001 T040301A (01         ) 00000                             2 YEARS OR LESS
002 T040301B (02         ) 10000                             3-5 YEARS
003 T040301C (03         ) 01000                             6-10 YEARS
004 T040301D (04         ) 00100                             11-24 YEARS
005 T040301E (05         ) 00010                             25 YEARS OR MORE
006 T040301M (M          ) 00001                             MISSING
```

```
CONDITIONING VARIABLE ID:  TCHR0006
DESCRIPTION:               TYPE TCHNG CERT IN THIS ST IN MAIN FIELD?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056201                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     7
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         6

001 T056201A (01         ) 000000                          ADVANCED PROFESSIONL
002 T056201B (02         ) 100000                          REGULAR/STANDARD ST
003 T056201C (03         ) 010000                          PROBATIONARY STATE
004 T056201D (04         ) 001000                          TEMPORARY/PROVISIONL
005 T056201E (05         ) 000100                          OTHER THAN STATE CRT
006 T056201F (06         ) 000010                          NOT HAVE CERT MAIN
007 T056201M (M          ) 000001                          MISSING

CONDITIONING VARIABLE ID:  TCHR0007
DESCRIPTION:               CERTIFICATION, ELEMENTARY OR MIDDLE/JUNIOR HS ED?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040501                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         3

001 T040501Y (01         ) 000                             YES
002 T040501N (02         ) 100                             NO
003 T040501C (03         ) 010                             NOT OFFERED IN STATE
004 T040501M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  TCHR0010
DESCRIPTION:               CERTIFICATION, ELEMENTARY SCIENCE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040507                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         3

001 T040507Y (01         ) 000                             YES
002 T040507N (02         ) 100                             NO
003 T040507C (03         ) 010                             NOT OFFERED IN STATE
004 T040507M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  TCHR0011
DESCRIPTION:               CERTIFICATION, MIDDLE/JUNIOR SCIENCE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040508                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         3

001 T040508Y (01         ) 000                             YES
002 T040508N (02         ) 100                             NO
003 T040508C (03         ) 010                             NOT OFFERED IN STATE
004 T040508M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  TCHR0012
DESCRIPTION:               CERTIFICATION, OTHER
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040505                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         3

001 T040505Y (01         ) 000                             YES
002 T040505N (02         ) 100                             NO
003 T040505C (03         ) 010                             NOT OFFERED IN STATE
004 T040505M (M          ) 001                             MISSING

CONDITIONING VARIABLE ID:  TCHR0013
DESCRIPTION:               HIGHEST ACADEMIC DEGREE YOU HOLD?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056301                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     8
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         7

001 T056301A (01         ) 0000000                         HIGH SCHOOL DIPLOMA
002 T056301B (02         ) 1000000                         ASSOCIATES/VOCATIONL
003 T056301C (03         ) 0100000                         BACHELORS DEGREE
004 T056301D (04         ) 0010000                         MASTER'S DEGREE
005 T056301E (05         ) 0001000                         EDUCATION SPECIALIST
006 T056301F (06         ) 0000100                         DOCTORATE
007 T056301G (07         ) 0000010                         PROFESSIONAL DEGREE
008 T056301M (M          ) 0000001                         MISSING

CONDITIONING VARIABLE ID:  TCHR0014
DESCRIPTION:               EDUCATION UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040701                         TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 T040701Y (01         ) 0                               YES
```

```
002 T040701M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0015
DESCRIPTION:               ELMENT ED UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040706                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040706Y (01         ) 0                                YES
002 T040706M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0016
DESCRIPTION:               SEC ED UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040707                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040707Y (01         ) 0                                YES
002 T040707M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0019
DESCRIPTION:               SCIENCE ED UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040710                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040710Y (01         ) 0                                YES
002 T040710M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0020
DESCRIPTION:               LIFE SCIENCE UNDERGRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040711                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040711Y (01         ) 0                                YES
002 T040711M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0021
DESCRIPTION:               PHYSICAL SCIENCE UNDERGRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040712                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040712Y (01         ) 0                                YES
002 T040712M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0022
DESCRIPTION:               EARTH SCIENCE UNDERGRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040713                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040713Y (01         ) 0                                YES
002 T040713M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0023
DESCRIPTION:               SPECIAL EDUCATION UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040708                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040708Y (01         ) 0                                YES
002 T040708M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0024
DESCRIPTION:               BILINGUAL ED/ESL  UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040709                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040709Y (01         ) 0                                YES
002 T040709M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0025
DESCRIPTION:               OTHER UNDERGRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040705                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
```

```
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040705Y (01         ) 0                                  YES
002 T040705M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0026
DESCRIPTION:               EDUCATION GRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040801                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040801Y (01         ) 0                                  YES
002 T040801M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0027
DESCRIPTION:               ELEMENTARY ED GRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040807                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040807Y (01         ) 0                                  YES
002 T040807M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0028
DESCRIPTION:               SECONDARY ED GRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040808                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040808Y (01         ) 0                                  YES
002 T040808M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0031
DESCRIPTION:               SCIENCE ED GRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040814                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040814Y (01         ) 0                                  YES
002 T040814M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0032
DESCRIPTION:               LIFE SCIENCE GRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040815                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040815Y (01         ) 0                                  YES
002 T040815M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0033
DESCRIPTION:               PHYSICAL SCIENCE GRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040816                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040816Y (01         ) 0                                  YES
002 T040816M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0034
DESCRIPTION:               EARTH SCIENCE GRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040817                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040817Y (01         ) 0                                  YES
002 T040817M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0035
DESCRIPTION:               SPECIAL ED GRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040809                            TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                              NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040809Y (01         ) 0                                  YES
002 T040809M (M          ) 1                                  MISSING

CONDITIONING VARIABLE ID:  TCHR0036
DESCRIPTION:               BILINGUAL GRAD MAJOR
```

```
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040810                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040810Y (01         ) 0                                YES
002 T040810M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0037
DESCRIPTION:               ADMIN/SUPERVISION GRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040811                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040811Y (01         ) 0                                YES
002 T040811M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0038
DESCRIPTION:               CURRICULUM/INSTRUCTION GRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040812                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040812Y (01         ) 0                                YES
002 T040812M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0039
DESCRIPTION:               COUNSELING GRAD MAJOR?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040813                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040813Y (01         ) 0                                YES
002 T040813M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0040
DESCRIPTION:               OTHER GRAD MAJOR
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040805                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040805Y (01         ) 0                                YES
002 T040805M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0041
DESCRIPTION:               NO GRADUATE STUDY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T040806                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T040806Y (01         ) 0                                YES
002 T040806M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0042
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-EDUCATION
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056401                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056401Y (01         ) 0                                YES
002 T056401M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0043
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-ELEMENTARY ED
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056402                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056402Y (01         ) 0                                YES
002 T056402M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0044
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-SECONDARY ED
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056403                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056403Y (01         ) 0                                YES
002 T056403M (M          ) 1                                MISSING
```

```
CONDITIONING VARIABLE ID:  TCHR0046
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-MATHEMATICS ED
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056405                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056405Y (01        ) 0                                 YES
002 T056405M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0047
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-SCIENCE ED
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056413                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056413Y (01        ) 0                                 YES
002 T056413M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0048
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-LIFE SCIENCE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056414                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056414Y (01        ) 0                                 YES
002 T056414M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0049
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-PHYSICAL SCIENCE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056415                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056415Y (01        ) 0                                 YES
002 T056415M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0051
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-SPECIAL ED
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056406                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056406Y (01        ) 0                                 YES
002 T056406M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0052
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-BILINGUAL ED
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056407                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056407Y (01        ) 0                                 YES
002 T056407M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0053
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-ADMIN & SUPERVISION
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056408                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056408Y (01        ) 0                                 YES
002 T056408M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0054
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-CURRICULUM & INSTRUC
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056409                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056409Y (01        ) 0                                 YES
002 T056409M (M         ) 1                                 MISSING

CONDITIONING VARIABLE ID:  TCHR0055
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-COUNSELING
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056410                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1
```

```
001 T056410Y (01         ) 0                               YES
002 T056410M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0056
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-OTHER
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056411                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056411Y (01         ) 0                               YES
002 T056411M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0057
DESCRIPTION:               UNDERGRAD/GRAD MINOR STUDY-NONE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056412                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056412Y (01         ) 0                               YES
002 T056412M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0062
DESCRIPTION:               PAST 5 YRS, TAKEN COURSES/IN PRO DEVP-TELECOMM USE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056701                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056701Y (01         ) 0                               YES
002 T056701M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0063
DESCRIPTION:               PAST 5 YRS, TAKEN COURSES/IN PRO DEVP-TECH USE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056702                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056702Y (01         ) 0                               YES
002 T056702M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0064
DESCRIPTION:               PAST 5 YRS, TAKEN COURSES/IN PRO DEVP-COOP INSTRCT
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056703                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056703Y (01         ) 0                               YES
002 T056703M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0065
DESCRIPTION:               PAST 5 YRS, COURSES/IN PRO DEVLP-INTERDISP INSTRCT
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056704                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056704Y (01         ) 0                               YES
002 T056704M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0066
DESCRIPTION:               PAST 5 YRS, COURSES/IN PRO DEVLP-PORTFOLIO ASSMNT
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056705                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056705Y (01         ) 0                               YES
002 T056705M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0067
DESCRIPTION:               PAST 5 YRS, COURSES/IN PRO DEVLP-PERF BASED ASSMNT
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T056706                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056706Y (01         ) 0                               YES
002 T056706M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0068
DESCRIPTION:               PAST 5 YRS, COURSES/PRO DEVLP-TEACH HIGHORDER THKG
GRADES/ASSESSMENTS:        N04, N08, S08
```

```
CONDITIONING VAR LABEL:
NAEP ID:                  T056707                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056707Y (01         ) 0                               YES
002 T056707M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0069
DESCRIPTION:              PAST 5 YRS, COURSES/PRO DEVLP-TEACH DIFF CULT BKGD
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056708                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056708Y (01         ) 0                               YES
002 T056708M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0070
DESCRIPTION:              PAST 5 YRS, COURSES/PRO DEVLP-TEACH LEP STUDENTS
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056709                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056709Y (01         ) 0                               YES
002 T056709M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0071
DESCRIPTION:              PAST 5 YRS, COURSES/PRO DEVLP-TEACH SPEC NEED STDS
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056710                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056710Y (01         ) 0                               YES
002 T056710M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0072
DESCRIPTION:              PAST 5 YRS, COURSES/PRO DEVLP-CLASSRM MNGMT/ORG
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056711                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056711Y (01         ) 0                               YES
002 T056711M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0073
DESCRIPTION:              PAST 5 YRS, COURSES/PRO DEVLP-OTHER PROF ISSUES
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056712                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056712Y (01         ) 0                               YES
002 T056712M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0074
DESCRIPTION:              PAST 5 YRS, COURSES/PRO DEVLP-NONE OF ABOVE
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056713                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T056713Y (01         ) 0                               YES
002 T056713M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TCHR0075
DESCRIPTION:              AVAILABILITY OF RESOURCES
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T041201                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 T041201A (01         ) 0000                            GET ALL RESOURCES
002 T041201B (02         ) 1000                            GET MOST RESOURCES
003 T041201C (03         ) 0100                            GET SOME RESOURCES
004 T041201D (04         ) 0010                            DONT GET RESOURCES
005 T041201M (M          ) 0001                            MISSING

CONDITIONING VARIABLE ID:  TCHR0078
DESCRIPTION:              HOW MANY SCHOOL HOURS ARE PREP TIME PER WEEK?
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T056801                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    7
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        6
```

```
001 T056801A (01         ) 000000                           NONE
002 T056801B (02         ) 100000                           LESS THAN ONE
003 T056801C (03         ) 010000                           1-2
004 T056801D (04         ) 001000                           3-4
005 T056801E (05         ) 000100                           5
006 T056801F (06         ) 000010                           MORE THAN 5
007 T056801M (M          ) 000001                           MISSING

CONDITIONING VARIABLE ID:  TCHR0080
DESCRIPTION:               METHODS OF TCHING SCI? COLLEGE COURSE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060301                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060301Y (01         ) 0                                YES
002 T060301M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0081
DESCRIPTION:               METHODS OF TCHING SCI?WRKSHP >1 WK
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060311                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060311Y (01         ) 0                                YES
002 T060311M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0082
DESCRIPTION:               METHODS OF TCHING SCI?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060321                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060321Y (01         ) 0                                YES
002 T060321M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0083
DESCRIPTION:               METHODS OF TCHING SCI?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060331                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060331Y (01         ) 0                                YES
002 T060331M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0084
DESCRIPTION:               METHODS OF TCHING SCI?OTHER PROF. DEV
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060341                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060341Y (01         ) 0                                YES
002 T060341M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0085
DESCRIPTION:               UNIV COURSES IN-BIO/LIFE SCI? COLLEGE COURSE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060302                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060302Y (01         ) 0                                YES
002 T060302M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0086
DESCRIPTION:               UNIV COURSES IN-BIO/LIFE SCI?WRKSHP >1 WK
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060312                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060312Y (01         ) 0                                YES
002 T060312M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0087
DESCRIPTION:               UNIV COURSES IN-BIO/LIFE SCI?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060322                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060322Y (01         ) 0                                YES
002 T060322M (M          ) 1                                MISSING
```

```
CONDITIONING VARIABLE ID:  TCHR0088
DESCRIPTION:               UNIV COURSES IN-BIO/LIFE SCI?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060332                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060332Y (01         ) 0                                YES
002 T060332M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0089
DESCRIPTION:               UNIV COURSES IN-BIO/LIFE SCI?OTHER PROF. DEV
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060342                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060342Y (01         ) 0                                YES
002 T060342M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0090
DESCRIPTION:               UNIV COURSES IN-CHEMISTRY? COLLEGE COURSE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060303                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060303Y (01         ) 0                                YES
002 T060303M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0091
DESCRIPTION:               UNIV COURSES IN-CHEMISTRY?WRKSHP >1 WK
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060313                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060313Y (01         ) 0                                YES
002 T060313M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0092
DESCRIPTION:               UNIV COURSES IN-CHEMISTRY?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060323                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060323Y (01         ) 0                                YES
002 T060323M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0093
DESCRIPTION:               UNIV COURSES IN-CHEMISTRY?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060333                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060333Y (01         ) 0                                YES
002 T060333M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0094
DESCRIPTION:               UNIV COURSES IN-CHEMISTRY?OTHER PROF. DEV
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060343                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060343Y (01         ) 0                                YES
002 T060343M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0095
DESCRIPTION:               UNIV COURSES IN-PHYSICS? COLLEGE COURSE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060304                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060304Y (01         ) 0                                YES
002 T060304M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0096
DESCRIPTION:               UNIV COURSES IN-PHYSICS?WRKSHP >1 WK
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060314                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1
```

```
001 T060314Y (01         ) 0                                YES
002 T060314M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0097
DESCRIPTION:               UNIV COURSES IN-PHYSICS?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060324                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060324Y (01         ) 0                                YES
002 T060324M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0098
DESCRIPTION:               UNIV COURSES IN-PHYSICS?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060334                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060334Y (01         ) 0                                YES
002 T060334M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0099
DESCRIPTION:               UNIV COURSES IN-PHYSICS?OTHER PROF. DEV
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060344                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060344Y (01         ) 0                                YES
002 T060344M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0100
DESCRIPTION:               UNIV COURSES IN-EARTH SCI? COLLEGE COURSE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060305                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060305Y (01         ) 0                                YES
002 T060305M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0101
DESCRIPTION:               UNIV COURSES IN-EARTH SCI?WRKSHP >1 WK
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060315                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060315Y (01         ) 0                                YES
002 T060315M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0102
DESCRIPTION:               UNIV COURSES IN-EARTH SCI?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060325                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060325Y (01         ) 0                                YES
002 T060325M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0103
DESCRIPTION:               UNIV COURSES IN-EARTH SCI?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060335                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060335Y (01         ) 0                                YES
002 T060335M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0104
DESCRIPTION:               UNIV COURSES IN-EARTH SCI?OTHER PROF. DEV
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060345                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 T060345Y (01         ) 0                                YES
002 T060345M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0105
DESCRIPTION:               UNIV COURSES-OTHER TYPES OF SCI? COLLEGE COURSE
GRADES/ASSESSMENTS:        N04, N08, S08
```

```
CONDITIONING VAR LABEL:
NAEP ID:                  T060306                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060306Y (01        ) 0                                YES
002 T060306M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0106
DESCRIPTION:              UNIV COURSES-OTHER TYPES OF SCI?WRKSHP >1 WK
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060316                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060316Y (01        ) 0                                YES
002 T060316M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0107
DESCRIPTION:              UNIV COURSES-OTHR TYPES OF SCI?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060326                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060326Y (01        ) 0                                YES
002 T060326M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0108
DESCRIPTION:              UNIV COURSES-OTHER TYPES OF SCI?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060336                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060336Y (01        ) 0                                YES
002 T060336M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0109
DESCRIPTION:              UNIV COURSES-OTHER TYPES OF SCI?OTHER PROF. DEV
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060346                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060346Y (01        ) 0                                YES
002 T060346M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0110
DESCRIPTION:              UNIV COURSES IN-NONE OF ABOVE? COLLEGE COURSE
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060307                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060307Y (01        ) 0                                YES
002 T060307M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0111
DESCRIPTION:              UNIV COURSES IN-NONE OF ABOVE?WRKSHP >1 WK
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060317                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060317Y (01        ) 0                                YES
002 T060317M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0112
DESCRIPTION:              UNIV COURSES IN-NONE OF ABOVE?WRKSHP <1 WK >1 DAY
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060327                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060327Y (01        ) 0                                YES
002 T060327M (M         ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0113
DESCRIPTION:              UNIV COURSES IN-NONE OF ABOVE?WRKSHP <= 1 DAY
GRADES/ASSESSMENTS:       N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060337                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060337Y (01        ) 0                                YES
002 T060337M (M         ) 1                                MISSING
```

```
CONDITIONING VARIABLE ID:  TCHR0114
DESCRIPTION:               UNIV COURSES IN-NONE OF ABOVE?OTHER PROF. DEV
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060347                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060347Y (01         ) 0                                YES
002 T060347M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0115
DESCRIPTION:               PAST 5 YRS, COURSES/ACTVTS IN-COMP USE TO GET DATA
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060401                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060401Y (01         ) 0                                YES
002 T060401M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0116
DESCRIPTION:               PAST 5 YRS, COURSES/ACTVTS IN-COMP DATA ANALYSIS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060402                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060402Y (01         ) 0                                YES
002 T060402M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0117
DESCRIPTION:               PAST 5 YRS, COURSES/ACTVTS IN-MULTIMEDIA SCI ED?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060403                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060403Y (01         ) 0                                YES
002 T060403M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0118
DESCRIPTION:               PAST 5 YRS, COURSES/ACTVTS IN-LAB MNGMT/SAFETY?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060404                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060404Y (01         ) 0                                YES
002 T060404M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0119
DESCRIPTION:               PAST 5 YRS, COURSES/ACTVTS IN-INTEGRATED SCI INST?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060405                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T060405Y (01         ) 0                                YES
002 T060405M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TCHR0120
DESCRIPTION:               YOU BELONG TO 1 OR > SCI RELATED SCI ORGS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060501                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 T060501Y (01         ) 00                               YES
002 T060501N (02         ) 10                               NO
003 T060501M (M          ) 01                               MISSING

CONDITIONING VARIABLE ID:  TSUB0001
DESCRIPTION:               HOW OFTEN STUDS READ SCI TEXTBOOK?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060601                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060601A (01         ) 0000                             ALMOST EVERY DAY
002 T060601B (02         ) 1000                             ONCE OR TWICE A WEEK
003 T060601C (03         ) 0100                             ONCE OR TWICE/MONTH
004 T060601D (04         ) 0010                             NEVER OR HARDLY EVER
005 T060601M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0002
DESCRIPTION:               HOW OFTEN STUDS READ BOOK/MAN ABOUT SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
```

```
CONDITIONING VAR LABEL:
NAEP ID:                  T060602                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060602A (01         ) 0000                           ALMOST EVERY DAY
002 T060602B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060602C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060602D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060602M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0003
DESCRIPTION:               HOW OFTEN STUDS DISCUSS SCI IN THE NEWS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060603                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060603A (01         ) 0000                           ALMOST EVERY DAY
002 T060603B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060603C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060603D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060603M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0004
DESCRIPTION:               HOW OFTEN STUDS WORK W/ OTHER STUDS ON ACT/PROJCT?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060604                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060604A (01         ) 0000                           ALMOST EVERY DAY
002 T060604B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060604C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060604D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060604M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0005
DESCRIPTION:               HOW OFTEN STUDS GIVE ORAL SCI REPORT?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060605                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060605A (01         ) 0000                           ALMOST EVERY DAY
002 T060605B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060605C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060605D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060605M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0006
DESCRIPTION:               HOW OFTEN STUDS PREPARE A WRITTEN SCI REPORT?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060606                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060606A (01         ) 0000                           ALMOST EVERY DAY
002 T060606B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060606C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060606D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060606M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0007
DESCRIPTION:               HOW OFTEN STUDS DO HANDS ON SCI ACTIVITIES IN SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060607                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060607A (01         ) 0000                           ALMOST EVERY DAY
002 T060607B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060607C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060607D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060607M (M          ) 0001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0008
DESCRIPTION:               HOW OFTEN STUDS TALK ABOUT MEASURES/RESULTS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  T060608                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:         CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060608A (01         ) 0000                           ALMOST EVERY DAY
002 T060608B (02         ) 1000                           ONCE OR TWICE A WEEK
003 T060608C (03         ) 0100                           ONCE OR TWICE/MONTH
004 T060608D (04         ) 0010                           NEVER OR HARDLY EVER
005 T060608M (M          ) 0001                           MISSING
```

```
CONDITIONING VARIABLE ID:  TSUB0009
DESCRIPTION:               HOW OFTEN STUDS TAKE SCI TEST OR QUIZ?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060609                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060609A (01         ) 0000                              ALMOST EVERY DAY
002 T060609B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060609C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060609D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060609M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0010
DESCRIPTION:               HOW OFTEN STUDS USE LIBRARY RESOURCES FOR SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060610                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060610A (01         ) 0000                              ALMOST EVERY DAY
002 T060610B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060610C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060610D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060610M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0011
DESCRIPTION:               HOW OFTEN STUDS USE COMPUTERS FOR SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060611                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060611A (01         ) 0000                              ALMOST EVERY DAY
002 T060611B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060611C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060611D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060611M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0012
DESCRIPTION:               HOW OFTEN DO YOU TALK TO CLASS ABOUT SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060701                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060701A (01         ) 0000                              ALMOST EVERY DAY
002 T060701B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060701C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060701D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060701M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0013
DESCRIPTION:               HOW OFTEN DO YOU DO A SCI DEMONSTRATION?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060702                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060702A (01         ) 0000                              ALMOST EVERY DAY
002 T060702B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060702C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060702D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060702M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0014
DESCRIPTION:               HOW OFTEN DO YOU SHOW A SCI VIDEOTAPE/TV PROGRAM?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060703                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060703A (01         ) 0000                              ALMOST EVERY DAY
002 T060703B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060703C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060703D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060703M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0015
DESCRIPTION:               HOW OFTEN DO YOU USE COMPUTERS FOR SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060704                           TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:        4

001 T060704A (01         ) 0000                              ALMOST EVERY DAY
```

```
002 T060704B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060704C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060704D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060704M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0016
DESCRIPTION:               HOW OFTEN DO YOU USE CDS OR LASER DISKS ON SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060705                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     5
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         4

001 T060705A (01         ) 0000                              ALMOST EVERY DAY
002 T060705B (02         ) 1000                              ONCE OR TWICE A WEEK
003 T060705C (03         ) 0100                              ONCE OR TWICE/MONTH
004 T060705D (04         ) 0010                              NEVER OR HARDLY EVER
005 T060705M (M          ) 0001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0017
DESCRIPTION:               HOW OFTEN YOUR SCI STUDS GO ON SCI FIELD TRIPS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060801                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         3

001 T060801A (01         ) 000                               3 OR MORE TIMES/YEAR
002 T060801B (02         ) 100                               1 OR 2 TIMES A YEAR
003 T060801C (03         ) 010                               NEVER OR HARDLY EVER
004 T060801M (M          ) 001                               MISSING

CONDITIONING VARIABLE ID:  TSUB0018
DESCRIPTION:               HOW OFTEN DO YOU BRING GUEST SPEAKER FOR SCI STUDS
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T060901                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         3

001 T060901A (01         ) 000                               3 OR MORE TIMES/YEAR
002 T060901B (02         ) 100                               1 OR 2 TIMES A YEAR
003 T060901C (03         ) 010                               NEVER OR HARDLY EVER
004 T060901M (M          ) 001                               MISSING

CONDITIONING VARIABLE ID:  TSUB0019
DESCRIPTION:               SAVE STUDS SCI WORK IN PORTFOLIOS FOR ASSESSMENT?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061001                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         2

001 T061001Y (01         ) 00                                YES
002 T061001N (02         ) 10                                NO
003 T061001M (M          ) 01                                MISSING

CONDITIONING VARIABLE ID:  TSUB0020
DESCRIPTION:               HOW MUCH EMPHASIS-KNOWING SCI FACTS/TERMS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061101                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         3

001 T061101A (01         ) 000                               HEAVY EMPHASIS
002 T061101B (02         ) 100                               MODERATE EMPHASIS
003 T061101C (03         ) 010                               LITTLE/NO EMPHASIS
004 T061101M (M          ) 001                               MISSING

CONDITIONING VARIABLE ID:  TSUB0021
DESCRIPTION:               HOW MUCH EMPHASIS-UNDERSTANDING KEY SCI CONCEPTS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061102                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         3

001 T061102A (01         ) 000                               HEAVY EMPHASIS
002 T061102B (02         ) 100                               MODERATE EMPHASIS
003 T061102C (03         ) 010                               LITTLE/NO EMPHASIS
004 T061102M (M          ) 001                               MISSING

CONDITIONING VARIABLE ID:  TSUB0022
DESCRIPTION:               HOW MUCH EMPHASIS-DEVELOP SCI PROB SOLVING SKILL?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061103                           TOTAL NUMBER OF SPECIFIED CONTRASTS:     4
TYPE OF CONTRAST:          CLASS                             NUMBER OF INDEPENDENT CONTRASTS:         3

001 T061103A (01         ) 000                               HEAVY EMPHASIS
002 T061103B (02         ) 100                               MODERATE EMPHASIS
003 T061103C (03         ) 010                               LITTLE/NO EMPHASIS
```

```
004 T061103M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0023
DESCRIPTION:               HOW MUCH EMPHASIS-SCI RELEVANCE TO SOCIETY/TECH?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061104                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 T061104A (01         ) 000                              HEAVY EMPHASIS
002 T061104B (02         ) 100                              MODERATE EMPHASIS
003 T061104C (03         ) 010                              LITTLE/NO EMPHASIS
004 T061104M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0024
DESCRIPTION:               HOW MUCH EMPHASIS-COMMUNICATE IDEAS IN SCI?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061105                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 T061105A (01         ) 000                              HEAVY EMPHASIS
002 T061105B (02         ) 100                              MODERATE EMPHASIS
003 T061105C (03         ) 010                              LITTLE/NO EMPHASIS
004 T061105M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0025
DESCRIPTION:               HOW MUCH EMPHASIS-DEVELOPING LAB SKILLS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061106                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 T061106A (01         ) 000                              HEAVY EMPHASIS
002 T061106B (02         ) 100                              MODERATE EMPHASIS
003 T061106C (03         ) 010                              LITTLE/NO EMPHASIS
004 T061106M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0026
DESCRIPTION:               HOW MUCH EMPHASIS-DEVELOPING STUDS SCI INTEREST?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061107                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 T061107A (01         ) 000                              HEAVY EMPHASIS
002 T061107B (02         ) 100                              MODERATE EMPHASIS
003 T061107C (03         ) 010                              LITTLE/NO EMPHASIS
004 T061107M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0027
DESCRIPTION:               HOW MUCH EMPHASIS-DEVELOPING DATA ANALYSIS SKILLS
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061108                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 T061108A (01         ) 000                              HEAVY EMPHASIS
002 T061108B (02         ) 100                              MODERATE EMPHASIS
003 T061108C (03         ) 010                              LITTLE/NO EMPHASIS
004 T061108M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0028
DESCRIPTION:               HOW MUCH EMPHASIS-USING TECH AS SCI TOOL?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061109                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 T061109A (01         ) 000                              HEAVY EMPHASIS
002 T061109B (02         ) 100                              MODERATE EMPHASIS
003 T061109C (03         ) 010                              LITTLE/NO EMPHASIS
004 T061109M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0029
DESCRIPTION:               EVER ASSIGN SOLO/GROUP SCI PROJECTS THAT TAKE >WK?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061201                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 T061201Y (01         ) 00                               YES
002 T061201N (02         ) 10                               NO
003 T061201M (M          ) 01                               MISSING

CONDITIONING VARIABLE ID:  TSUB0030
DESCRIPTION:               HOW OFTEN USE MULT CHOICE TESTS TO ACCESS?
```

```
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061301                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:           CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061301A (01         ) 00000                             ONCE OR TWICE WEEK
002 T061301B (02         ) 10000                             ONCE OR TWICE MONTH
003 T061301C (03         ) 01000                             ONCE/GRADING PERIOD
004 T061301D (04         ) 00100                             ONCE OR TWICE/YEAR
005 T061301E (05         ) 00010                             NEVER OR HARDLY EVER
006 T061301M (M          ) 00001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0031
DESCRIPTION:                HOW OFTEN USE SHOR/LONG WRITTEN RESPONSE TO ACCESS
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061302                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:           CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061302A (01         ) 00000                             ONCE OR TWICE WEEK
002 T061302B (02         ) 10000                             ONCE OR TWICE MONTH
003 T061302C (03         ) 01000                             ONCE/GRADING PERIOD
004 T061302D (04         ) 00100                             ONCE OR TWICE/YEAR
005 T061302E (05         ) 00010                             NEVER OR HARDLY EVER
006 T061302M (M          ) 00001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0032
DESCRIPTION:                HOW OFTEN USE SOLO PROJECTS TO ACCESS?
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061303                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:           CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061303A (01         ) 00000                             ONCE OR TWICE WEEK
002 T061303B (02         ) 10000                             ONCE OR TWICE MONTH
003 T061303C (03         ) 01000                             ONCE/GRADING PERIOD
004 T061303D (04         ) 00100                             ONCE OR TWICE/YEAR
005 T061303E (05         ) 00010                             NEVER OR HARDLY EVER
006 T061303M (M          ) 00001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0033
DESCRIPTION:                HOW OFTEN USE GROUP PROJECTS TO ACCESS?
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061304                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:           CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061304A (01         ) 00000                             ONCE OR TWICE WEEK
002 T061304B (02         ) 10000                             ONCE OR TWICE MONTH
003 T061304C (03         ) 01000                             ONCE/GRADING PERIOD
004 T061304D (04         ) 00100                             ONCE OR TWICE/YEAR
005 T061304E (05         ) 00010                             NEVER OR HARDLY EVER
006 T061304M (M          ) 00001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0034
DESCRIPTION:                HOW OFTEN USE WORK PORTFOLIOS TO ACCESS?
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061305                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:           CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061305A (01         ) 00000                             ONCE OR TWICE WEEK
002 T061305B (02         ) 10000                             ONCE OR TWICE MONTH
003 T061305C (03         ) 01000                             ONCE/GRADING PERIOD
004 T061305D (04         ) 00100                             ONCE OR TWICE/YEAR
005 T061305E (05         ) 00010                             NEVER OR HARDLY EVER
006 T061305M (M          ) 00001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0035
DESCRIPTION:                HOW OFTEN USE IN CLASS ESSAYS TO ACCESS?
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061306                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:           CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061306A (01         ) 00000                             ONCE OR TWICE WEEK
002 T061306B (02         ) 10000                             ONCE OR TWICE MONTH
003 T061306C (03         ) 01000                             ONCE/GRADING PERIOD
004 T061306D (04         ) 00100                             ONCE OR TWICE/YEAR
005 T061306E (05         ) 00010                             NEVER OR HARDLY EVER
006 T061306M (M          ) 00001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0036
DESCRIPTION:                HOW OFTEN USE SELF/PEER EVAL TO ACCESS?
GRADES/ASSESSMENTS:         N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                    T061307                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
```

```
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061307A (01         ) 00000                                 ONCE OR TWICE WEEK
002 T061307B (02         ) 10000                                 ONCE OR TWICE MONTH
003 T061307C (03         ) 01000                                 ONCE/GRADING PERIOD
004 T061307D (04         ) 00100                                 ONCE OR TWICE/YEAR
005 T061307E (05         ) 00010                                 NEVER OR HARDLY EVER
006 T061307M (M          ) 00001                                 MISSING

CONDITIONING VARIABLE ID:  TSUB0037
DESCRIPTION:               HOW OFTEN USE LAB NOTEBOOKS/JOURNALS TO ACCESS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061308                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061308A (01         ) 00000                                 ONCE OR TWICE WEEK
002 T061308B (02         ) 10000                                 ONCE OR TWICE MONTH
003 T061308C (03         ) 01000                                 ONCE/GRADING PERIOD
004 T061308D (04         ) 00100                                 ONCE OR TWICE/YEAR
005 T061308E (05         ) 00010                                 NEVER OR HARDLY EVER
006 T061308M (M          ) 00001                                 MISSING

CONDITIONING VARIABLE ID:  TSUB0038
DESCRIPTION:               HOW OFTEN USE HOMEWORK TO ACCESS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061309                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061309A (01         ) 00000                                 ONCE OR TWICE WEEK
002 T061309B (02         ) 10000                                 ONCE OR TWICE MONTH
003 T061309C (03         ) 01000                                 ONCE/GRADING PERIOD
004 T061309D (04         ) 00100                                 ONCE OR TWICE/YEAR
005 T061309E (05         ) 00010                                 NEVER OR HARDLY EVER
006 T061309M (M          ) 00001                                 MISSING

CONDITIONING VARIABLE ID:  TSUB0039
DESCRIPTION:               HOW OFTEN USE HANDS ON ACTIVITIES TO ACCESS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061310                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061310A (01         ) 00000                                 ONCE OR TWICE WEEK
002 T061310B (02         ) 10000                                 ONCE OR TWICE MONTH
003 T061310C (03         ) 01000                                 ONCE/GRADING PERIOD
004 T061310D (04         ) 00100                                 ONCE OR TWICE/YEAR
005 T061310E (05         ) 00010                                 NEVER OR HARDLY EVER
006 T061310M (M          ) 00001                                 MISSING

CONDITIONING VARIABLE ID:  TSUB0040
DESCRIPTION:               PROPORTION OF EVAL IN SCI BASED ON HANDS ON ACTVS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061401                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        4

001 T061401A (01         ) 0000                                  MOST OR ALL OF GRADE
002 T061401B (02         ) 1000                                  ABOUT HALF OF GRADE
003 T061401C (03         ) 0100                                  VERY LITTLE OF GRADE
004 T061401D (04         ) 0010                                  NONE OF GRADE
005 T061401M (M          ) 0001                                  MISSING

CONDITIONING VARIABLE ID:  TSUB0041
DESCRIPTION:               BEST DESCRIPTION OF COMPUTER AVAILABILITY FOR SCI
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061501                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    7
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        6

001 T061501A (01         ) 000000                                NONE AVAILABLE
002 T061501B (02         ) 100000                                1 IN CLASSROOM
003 T061501C (03         ) 010000                                2-3 IN CLASSROOM
004 T061501D (04         ) 001000                                4 IN CLASSROOM
005 T061501E (05         ) 000100                                IN LAB BUT DIFFICULT
006 T061501F (06         ) 000010                                IN LAB AND EASY
007 T061501M (M          ) 000001                                MISSING

CONDITIONING VARIABLE ID:  TSUB0042
DESCRIPTION:               USE COMPUTERS FOR SCI INSTRUCTION: PLAYING SCI
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061601                               TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                                 NUMBER OF INDEPENDENT CONTRASTS:        1

001 T061601Y (01         ) 0                                     YES
```

```
002 T061601M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0043
DESCRIPTION:               USE COMPUTERS FOR SCI INSTRUCTION: PLAYING SCI
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061611                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T061611Y (01         ) 0                                YES
002 T061611M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0044
DESCRIPTION:               USE COMPUTERS FOR SCI INSTRUCTION: SIMULATIONS
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061621                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T061621Y (01         ) 0                                YES
002 T061621M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0045
DESCRIPTION:               USE COMPUTERS FOR SCI INSTRUCTION: DATA ANALYSIS
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061631                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T061631Y (01         ) 0                                YES
002 T061631M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0046
DESCRIPTION:               USE COMPUTERS FOR SCI INSTRUCTION: WORD PROCESS
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061641                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T061641Y (01         ) 0                                YES
002 T061641M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0047
DESCRIPTION:               USE COMPUTERS FOR SCI INSTRUCTION: DO NOT USE
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061651                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T061651Y (01         ) 0                                YES
002 T061651M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0048
DESCRIPTION:               STUDS ASSIGNED TO CLASS BY ABILITY/ACHVMNT LEVEL?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061701                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061701A (01         ) 00000                            NOT ASSIGNED BY ABIL
002 T061701B (02         ) 10000                            ASSIGNED BY MATN ABL
003 T061701C (03         ) 01000                            ASSIGNED BY READING
004 T061701D (04         ) 00100                            ASSIGNED BY SCIENCE
005 T061701E (05         ) 00010                            ASSIGNED BY GENERAL
006 T061701M (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  TSUB0049
DESCRIPTION:               IF ASSIGNED BY ABILITY WHICH BEST DESCRIBES LEVEL?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061801                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T061801A (01         ) 00000                            NOT ASSIGNED BY ABIL
002 T061801B (02         ) 10000                            MOSTLY HIGH ABILITY
003 T061801C (03         ) 01000                            MOSTLY AVERAGE ABIL
004 T061801D (04         ) 00100                            MOSTLY LOW ABILITY
005 T061801E (05         ) 00010                            MOSTLY MIXED ABILITY
006 T061801M (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  TSUB0050
DESCRIPTION:               COMPOSITION OF CLASS ACCORDING TO GENDER?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T061901                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3
```

```
001 T061901A (01         ) 000                              MIXED
002 T061901B (02         ) 100                              ALL MALE
003 T061901C (03         ) 010                              ALL FEMALE
004 T061901M (M          ) 001                              MISSING

CONDITIONING VARIABLE ID:  TSUB0051
DESCRIPTION:               HOW MUCH TIME CLASS SPEND ON LIFE SCIENCE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062001                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 T062001A (01         ) 0000                             A LOT
002 T062001B (02         ) 1000                             SOME
003 T062001C (03         ) 0100                             LITTLE
004 T062001D (04         ) 0010                             NONE
005 T062001M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0052
DESCRIPTION:               HOW MUCH TIME CLASS SPEND ON EARTH SCIENCE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062002                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 T062002A (01         ) 0000                             A LOT
002 T062002B (02         ) 1000                             SOME
003 T062002C (03         ) 0100                             LITTLE
004 T062002D (04         ) 0010                             NONE
005 T062002M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0053
DESCRIPTION:               HOW MUCH TIME CLASS SPEND ON PHYSICAL SCIENCE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062003                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 T062003A (01         ) 0000                             A LOT
002 T062003B (02         ) 1000                             SOME
003 T062003C (03         ) 0100                             LITTLE
004 T062003D (04         ) 0010                             NONE
005 T062003M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0054
DESCRIPTION:               WHICH BEST DESCRIBES SPACE WHERE CLASS TAUGHT?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062101                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    5
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        4

001 T062101A (01         ) 0000                             CLASS ONLY
002 T062101B (02         ) 1000                             CLASS W/ H2O ACCESS
003 T062101C (03         ) 0100                             CLASS W/ LAB ACCESS
004 T062101D (04         ) 0010                             LAB W/ H2O ACCESS
005 T062101M (M          ) 0001                             MISSING

CONDITIONING VARIABLE ID:  TSUB0055
DESCRIPTION:               DO STUDS PRODUCE NOTEBOOKS/REPORTS OF LAB WORK?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062201                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062201Y (01         ) 0                                YES
002 T062201M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0056
DESCRIPTION:               DO STUDS PRODUCE REPORTS OF EXTENDED SCI PROJECTS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062202                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062202Y (01         ) 0                                YES
002 T062202M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  TSUB0057
DESCRIPTION:               DO STUDS PRODUCE REPORTS ON SPECIFIC TOPIC/ISSUE?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062203                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062203Y (01         ) 0                                YES
002 T062203M (M          ) 1                                MISSING
```

```
CONDITIONING VARIABLE ID:  TSUB0058
DESCRIPTION:               DO STUDS PRODUCE REPORTS/RECORDS OF FIELD TRIPS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062204                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062204Y (01         ) 0                               YES
002 T062204M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0059
DESCRIPTION:               DO STUDS PRODUCE JOURNALS/DIARIES/LOGS OF IDEAS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062205                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062205Y (01         ) 0                               YES
002 T062205M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0060
DESCRIPTION:               DO STUDS PRODUCE PHOTO RECORDS OF PROJECTS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062206                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062206Y (01         ) 0                               YES
002 T062206M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0061
DESCRIPTION:               DO STUDS PRODUCE AUDIO/VIDEOTAPE RECORDS OF ACTVS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062207                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062207Y (01         ) 0                               YES
002 T062207M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0062
DESCRIPTION:               DO STUDS PRODUCE REPORTS OF PERSONAL INTERVIEWS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062208                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062208Y (01         ) 0                               YES
002 T062208M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0063
DESCRIPTION:               DO STUDS PRODUCE 3D SCI MODELS?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062209                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062209Y (01         ) 0                               YES
002 T062209M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0064
DESCRIPTION:               DO STUDS PRODUCE COMP GENERATED MULTMEDIA PROJECTS
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062210                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 T062210Y (01         ) 0                               YES
002 T062210M (M          ) 1                               MISSING

CONDITIONING VARIABLE ID:  TSUB0065
DESCRIPTION:               TIME PER WEEK EXPECT STUD TO SPEND ON HOMEWORK?
GRADES/ASSESSMENTS:        N04, N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062301                         TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                           NUMBER OF INDEPENDENT CONTRASTS:        5

001 T062301A (01         ) 00000                           NONE
002 T062301B (02         ) 10000                           1/2 HOUR
003 T062301C (03         ) 01000                           1 HOUR
004 T062301D (04         ) 00100                           2 HOURS
005 T062301E (05         ) 00010                           MORE THAN 2 HOURS
006 T062301M (M          ) 00001                           MISSING

CONDITIONING VARIABLE ID:  TSUB0066
DESCRIPTION:               CLASS PERIOD AND # OF STUDS IN CLASS
GRADES/ASSESSMENTS:        N04, N08, S08
```

```
CONDITIONING VAR LABEL:
NAEP ID:                  T062401                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T062401A (01        ) 00000                            1-20 STUDENTS
002 T062401B (02        ) 10000                            21025 STUDENTS
003 T062401C (03        ) 01000                            26-30 STUDENTS
004 T062401D (04        ) 00100                            31-35 STUDENTS
005 T062401E (05        ) 00010                            36 OR MORE STUDENTS
006 T062401M (M         ) 00001                            MISSING

CONDITIONING VARIABLE ID:  SCHL0087
DESCRIPTION:              BEST DESCRIBES HOW 8TH GRADES ARE ORGANIZED?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034201                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    4
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        3

001 C034201A (01        ) 000                              SELF-CONTAINED
002 C034201B (02        ) 100                              SEMI-DEPARTMENTALIZE
003 C034201C (03        ) 010                              DEPARTMENTALIZED
004 C034201M (M         ) 001                              MISSING

CONDITIONING VARIABLE ID:  SCHL0088
DESCRIPTION:              ARE 8TH-GRADERS ASSIGNED TO MATH BY ABILITY?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034402                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 C034402Y (01        ) 00                               YES
002 C034402N (02        ) 10                               NO
003 C034402M (M         ) 01                               MISSING

CONDITIONING VARIABLE ID:  SCHL0089
DESCRIPTION:              ARE 8TH-GRADERS ASSIGNED TO SCIENCE BY ABILITY?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034403                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 C034403Y (01        ) 00                               YES
002 C034403N (02        ) 10                               NO
003 C034403M (M         ) 01                               MISSING

CONDITIONING VARIABLE ID:  SCHL0090
DESCRIPTION:              ARE 8TH-GRADERS ASSIGNED TO ENGLISH BY ABILITY?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034401                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 C034401Y (01        ) 00                               YES
002 C034401N (02        ) 10                               NO
003 C034401M (M         ) 01                               MISSING

CONDITIONING VARIABLE ID:  SCHL0091
DESCRIPTION:              ARE 8TH-GRADERS ASSIGNED TO ARTS BY ABILITY?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034406                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 C034406Y (01        ) 00                               YES
002 C034406N (02        ) 10                               NO
003 C034406M (M         ) 01                               MISSING

CONDITIONING VARIABLE ID:  SCHL0092
DESCRIPTION:              HOW OFTEN 8TH GRDS RECEIVE COMP SCI INSTRUCTION?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034510                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 C034510A (01        ) 00000                            EVERY DAY
002 C034510B (02        ) 10000                            3 OR 4 TIMES A WEEK
003 C034510C (03        ) 01000                            ONCE OR TWICE/WEEK
004 C034510D (04        ) 00100                            LESS THAN ONCE/WEEK
005 C034510E (05        ) 00010                            SUBJECT NOT TAUGHT
006 C034510M (M         ) 00001                            MISSING

CONDITIONING VARIABLE ID:  SCHL0093
DESCRIPTION:              HOW OFTEN 8TH GRDS RECEIVE MATH INSTRUCTION?
GRADES/ASSESSMENTS:       N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                  C034511                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:         CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5
```

```
001 C034511A (01         ) 00000                          EVERY DAY
002 C034511B (02         ) 10000                          3 OR 4 TIMES A WEEK
003 C034511C (03         ) 01000                          ONCE OR TWICE/WEEK
004 C034511D (04         ) 00100                          LESS THAN ONCE/WEEK
005 C034511E (05         ) 00010                          SUBJECT NOT TAUGHT
006 C034511M (M          ) 00001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0094
DESCRIPTION:               HOW OFTEN 8TH GRDS RECEIVE SCIENCE INSTRUCTION?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C034512                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     6
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         5

001 C034512A (01         ) 00000                          EVERY DAY
002 C034512B (02         ) 10000                          3 OR 4 TIMES A WEEK
003 C034512C (03         ) 01000                          ONCE OR TWICE/WEEK
004 C034512D (04         ) 00100                          LESS THAN ONCE/WEEK
005 C034512E (05         ) 00010                          SUBJECT NOT TAUGHT
006 C034512M (M          ) 00001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0095
DESCRIPTION:               HOW OFTEN 8TH GRDS RECEIVE ENGLISH INSTRUCTION?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C034513                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     6
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         5

001 C034513A (01         ) 00000                          EVERY DAY
002 C034513B (02         ) 10000                          3 OR 4 TIMES A WEEK
003 C034513C (03         ) 01000                          ONCE OR TWICE/WEEK
004 C034513D (04         ) 00100                          LESS THAN ONCE/WEEK
005 C034513E (05         ) 00010                          SUBJECT NOT TAUGHT
006 C034513M (M          ) 00001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0096
DESCRIPTION:               HOW OFTEN 8TH GRDS RECEIVE ARTS INSTRUCTION?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C034514                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     6
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         5

001 C034514A (01         ) 00000                          EVERY DAY
002 C034514B (02         ) 10000                          3 OR 4 TIMES A WEEK
003 C034514C (03         ) 01000                          ONCE OR TWICE/WEEK
004 C034514D (04         ) 00100                          LESS THAN ONCE/WEEK
005 C034514E (05         ) 00010                          SUBJECT NOT TAUGHT
006 C034514M (M          ) 00001                          MISSING

CONDITIONING VARIABLE ID:  SCHL0097
DESCRIPTION:               HAS ENGLISH BEEN IDENTIFIED AS A PRIORITY?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C031611                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         2

001 C031611Y (01         ) 00                             YES
002 C031611N (02         ) 10                             NO
003 C031611M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0098
DESCRIPTION:               SCHOOL OFFER 8TH GR STUDS ALGEBRA FOR HS CREDIT?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C034601                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     3
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         2

001 C034601Y (01         ) 00                             YES
002 C034601N (02         ) 10                             NO
003 C034601M (M          ) 01                             MISSING

CONDITIONING VARIABLE ID:  SCHL0099
DESCRIPTION:               SCHOOL W/ SPECIAL FOCUS ON ENGLISH?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037203                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                          NUMBER OF INDEPENDENT CONTRASTS:         1

001 C037203Y (01         ) 0                              YES
002 C037203M (M          ) 1                              MISSING

CONDITIONING VARIABLE ID:  SCHL0100
DESCRIPTION:               SCHOOL FOLLOW DISTRICT/STATE ENGLISH CURRICULUM?
GRADES/ASSESSMENTS:        N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   C037306                        TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
```

```
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C037306Y (01         ) 0                                YES
002 C037306M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0101
DESCRIPTION:               SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR MATH?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039401                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039401Y (01         ) 0                                YES
002 C039401M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0102
DESCRIPTION:               SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR SCIENCE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039402                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039402Y (01         ) 0                                YES
002 C039402M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0103
DESCRIPTION:               SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR READING?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039403                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039403Y (01         ) 0                                YES
002 C039403M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0104
DESCRIPTION:               SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR ARTS?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039404                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039404Y (01         ) 0                                YES
002 C039404M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0105
DESCRIPTION:               SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR OTHER?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039405                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039405Y (01         ) 0                                YES
002 C039405M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0106
DESCRIPTION:               SCHOOL SPONSER 8TH GRDS FIELD TRIP FOR NONE ABOVE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039406                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039406Y (01         ) 0                                YES
002 C039406M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0107
DESCRIPTION:               8TH GRADERS IN EXTRACURR ACTS FOR MATH?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039501                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039501Y (01         ) 0                                YES
002 C039501M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0108
DESCRIPTION:               8TH GRADERS IN EXTRACURR ACTS FOR SCIENCE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039502                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 C039502Y (01         ) 0                                YES
002 C039502M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0109
DESCRIPTION:               8TH GRADERS IN EXTRACURR ACTS FOR ENG/LANG ARTS?
```

```
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039503                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039503Y (01         ) 0                                YES
002 C039503M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0110
DESCRIPTION:               8TH GRADERS IN EXTRACURR ACTS FOR ARTS?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039504                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039504Y (01         ) 0                                YES
002 C039504M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0111
DESCRIPTION:               8TH GRADERS IN EXTRACURR ACTS FOR NONE OF ABOVE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039505                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039505Y (01         ) 0                                YES
002 C039505M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0112
DESCRIPTION:               8TH GRADERS IN SUMMER PROGRAMS IN MATH?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039601                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039601Y (01         ) 0                                YES
002 C039601M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0113
DESCRIPTION:               8TH GRADERS IN SUMMER PROGRAMS IN SCIENCE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039602                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039602Y (01         ) 0                                YES
002 C039602M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0114
DESCRIPTION:               8TH GRADERS IN SUMMER PROGRAMS IN ENG/LANG ARTS?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039603                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039603Y (01         ) 0                                YES
002 C039603M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0115
DESCRIPTION:               8TH GRADERS IN SUMMER PROGRAMS IN ARTS?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039604                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039604Y (01         ) 0                                YES
002 C039604M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0116
DESCRIPTION:               8TH GRADERS IN SUMMER PROGRAMS IN NONE OF ABOVE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C039605                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 C039605Y (01         ) 0                                YES
002 C039605M (M          ) 1                                MISSING

CONDITIONING VARIABLE ID:  SCHL0117
DESCRIPTION:               WHAT % OF 8TH GRDS HELD BACK/REPEAT 8TH GRADE?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   C041901                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         5

001 C041901A (01         ) 00000                            0%
002 C041901B (02         ) 10000                            1-2%
```

```
003 C041901C (03         ) 01000                            3-5%
004 C041901D (04         ) 00100                            6-10%
005 C041901E (05         ) 00010                            MORE THAN 10%
006 C041901M (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  TCHR0121
DESCRIPTION:               COUNTING THIS YR, HOW MANY YRS TOTAL TAUGHT SCI?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062501                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T062501A (01         ) 00000                            2 YEARS OR LESS
002 T062501B (02         ) 10000                            3-5 YEARS
003 T062501C (03         ) 01000                            6-10 YEARS
004 T062501D (04         ) 00100                            11-24 YEARS
005 T062501E (05         ) 00010                            25 YEARS OR MORE
006 T062501M (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  TCHR0122
DESCRIPTION:               LAST YR, TIME IN PRO WORKSHOPS/SEMS IN SCI?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062601                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T062601A (01         ) 00000                            NONE
002 T062601B (02         ) 10000                            LESS THAN 6 HOURS
003 T062601C (03         ) 01000                            6-15 HOURS
004 T062601D (04         ) 00100                            16-35 HOURS
005 T062601E (05         ) 00010                            MORE THAN 35 HOURS
006 T062601M (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  TCHR0123
DESCRIPTION:               LAST 2 YRS, # OF UNIV COURSES IN SCI/SCI ED?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062701                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    6
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        5

001 T062701A (01         ) 00000                            NONE
002 T062701B (02         ) 10000                            ONE
003 T062701C (03         ) 01000                            TWO
004 T062701D (04         ) 00100                            THREE
005 T062701E (05         ) 00010                            FOUR OR MORE
006 T062701M (M          ) 00001                            MISSING

CONDITIONING VARIABLE ID:  TCHR0124
DESCRIPTION:               CURRICULUM SPECIALIST TO HELP/ADVISE IN SCI?
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   T062801                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    3
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        2

001 T062801Y (01         ) 00                               YES
002 T062801N (02         ) 10                               NO
003 T062801M (M          ) 01                               MISSING

CONDITIONING VARIABLE ID:  SCHL0173
DESCRIPTION:               PERCENT OF STUDENTS WHO PARTICIPATED IN THE NATIONAL SCHOOL LUNCH PROGRAM
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   NTLUNSC                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 NATLUNCM (M          ) 0                                MISSING
002 NATLUNCH (@          ) 1                                PERCENT

CONDITIONING VARIABLE ID:  SCHL0174
DESCRIPTION:               PERCENT OF STUDENTS WHO PARTICIPATED IN THE NATIONAL SCHOOL LUNCH PROGRAM
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   NTLUNSC                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    1
TYPE OF CONTRAST:          LINEAR                           NUMBER OF INDEPENDENT CONTRASTS:        1

001 NATLUNCL (0-100,M=0  )  0.0 +  1.0*X                    LINEAR

CONDITIONING VARIABLE ID:  SCHL0175
DESCRIPTION:               PERCENT OF STUDENTS WHO PARTICIPATED IN REMEDIAL READING
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   REMRDSC                          TOTAL NUMBER OF SPECIFIED CONTRASTS:    2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:        1

001 REMREADM (M          ) 0                                MISSING
002 REM READ (@          ) 1                                PERCENT
```

```
CONDITIONING VARIABLE ID:  SCHL0176
DESCRIPTION:               PERCENT OF STUDENTS WHO PARTICIPATED IN REMEDIAL READING
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   REMRDSC                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          LINEAR                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMREADL (0-100,M=0  )  0.0 +  1.0*X                    LINEAR

CONDITIONING VARIABLE ID:  SCHL0177
DESCRIPTION:               PERCENT OF STUDENTS WHO PARTICIPATED IN REMEDIAL MATH
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   REMMHSC                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMMATHM (M          ) 0                                MISSING
002 REM MATH (@          ) 1                                PERCENT

CONDITIONING VARIABLE ID:  SCHL0178
DESCRIPTION:               PERCENT OF STUDENTS WHO PARTICIPATED IN REMEDIAL MATH
GRADES/ASSESSMENTS:        N04, S04, N08, S08, N12
CONDITIONING VAR LABEL:
NAEP ID:                   REMMHSC                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          LINEAR                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMMATHL (0-100,M=0  )  0.0 +  1.0*X                    LINEAR

CONDITIONING VARIABLE ID:  SCHL0185
DESCRIPTION:               PERCENT OF STUDENTS IN 8TH WHO PARTICIPATED IN THE NATIONAL SCHOOL LUNCH PROGRAM
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   NTLUNGR                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 NATLUN8M (M          ) 0                                MISSING
002 NATLUN8  (@          ) 1                                PERCENT

CONDITIONING VARIABLE ID:  SCHL0186
DESCRIPTION:               PERCENT OF STUDENTS IN 8TH WHO PARTICIPATED IN THE NATIONAL SCHOOL LUNCH PROGRAM
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   NTLUNGR                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          LINEAR                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 NATLUN8L (0-100,M=0  )  0.0 +  1.0*X                    LINEAR

CONDITIONING VARIABLE ID:  SCHL0187
DESCRIPTION:               PERCENT OF STUDENTS IN 8TH WHO PARTICIPATED IN REMEDIAL READING
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   REMRDGR                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMREA8M (M          ) 0                                MISSING
002 REMREAD8 (@          ) 1                                PERCENT

CONDITIONING VARIABLE ID:  SCHL0188
DESCRIPTION:               PERCENT OF STUDENTS IN 8TH WHO PARTICIPATED IN REMEDIAL READING
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   REMRDGR                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          LINEAR                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMREA8L (0-100,M=0  )  0.0 +  1.0*X                    LINEAR

CONDITIONING VARIABLE ID:  SCHL0189
DESCRIPTION:               PERCENT OF STUDENTS IN 8TH WHO PARTICIPATED IN REMEDIAL MATH
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   REMMHGR                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     2
TYPE OF CONTRAST:          CLASS                            NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMMAT8M (M          ) 0                                MISSING
002 REMMATH8 (@          ) 1                                PERCENT

CONDITIONING VARIABLE ID:  SCHL0190
DESCRIPTION:               PERCENT OF STUDENTS IN 8TH WHO PARTICIPATED IN REMEDIAL MATH
GRADES/ASSESSMENTS:        N08, S08
CONDITIONING VAR LABEL:
NAEP ID:                   REMMHGR                          TOTAL NUMBER OF SPECIFIED CONTRASTS:     1
TYPE OF CONTRAST:          LINEAR                           NUMBER OF INDEPENDENT CONTRASTS:         1

001 REMMAT8L (0-100,M=0  )  0.0 +  1.0*X                    LINEAR
```

Appendix D

IRT PARAMETERS FOR SCIENCE ITEMS

This appendix contains three tables of item response theory (IRT) parameters for the items that were used in each field of science scale for the eighth-grade State Assessment.

For each of the binary scored items used in scaling (i.e., multiple-choice items and short constructed-response items), the tables provide estimates of the IRT parameters (which correspond to a_j, b_j, and c_j in Equation 8.1 in Chapter 8) and their associated standard errors (s.e.) of the estimates. For each of the polytomously scored items (i.e., the extended constructed-response items and the testlets), the tables also show the estimates of the d_{jv} parameters (see Equation 8.1) and their associated standard errors.

The tables also show the block in which each item appears (*Block*) and the position of each item within its block (*Item*).

Note that because the item parameters in this appendix are in the metrics used for the original calibration of the scales, the grade 8 parameters are shown in different metrics. The transformations needed to represent these parameters in terms of the metric of the final reporting scales are given in Chapter 9.

Table D-1

IRT Parameters for Science Items

Earth Science, Grade 8

NAEP ID	Block	Item	a_j (s.e.)	b_j (s.e.)	c_j (s.e.)	d_{j1} (s.e.)	d_{j2} (s.e.)	d_{j3} (s.e.)	d_{j4} (s.e.)	d_{j5} (s.e.)
2K040701	SD	1A	0.381 (0.009)	1.084 (0.029)	0.000 (0.000)	1.351 (0.044)	-3.323 (0.114)	1.972 (0.128)		
2K040706	SD	7	0.559 (0.040)	-2.203 (0.234)	0.466 (0.056)					
2K040708	SD	9A	0.547 (0.015)	1.418 (0.035)	0.000 (0.000)	-0.381 (0.041)	0.381 (0.056)			
2K040709	SD	10	0.516 (0.046)	-0.514 (0.199)	0.512 (0.038)					
2K0407CL	SD		0.765 (0.016)	-0.082 (0.013)	0.000 (0.000)	0.979 (0.032)	-0.719 (0.032)	-0.260 (0.033)		
2K040713	SD	14A	0.594 (0.013)	0.896 (0.017)	0.000 (0.000)	-0.535 (0.043)	-0.177 (0.062)	0.713 (0.060)		
2K040801	SE	0A	0.418 (0.013)	-1.116 (0.039)	0.000 (0.000)	0.039 (0.061)	-0.039 (0.045)			
2K040802	SE	1A	0.452 (0.012)	-1.883 (0.051)	0.000 (0.000)	-1.210 (0.088)	1.210 (0.065)			
2K040808	SE	1G	0.596 (0.012)	0.002 (0.018)	0.000 (0.000)	-1.344 (0.052)	1.344 (0.051)			
2K040809	SE	1M	0.608 (0.013)	0.558 (0.019)	0.000 (0.000)	-1.417 (0.054)	1.417 (0.056)			
2K040803	SE	2A	0.639 (0.024)	-0.224 (0.031)	0.000 (0.000)					
2K040804	SE	3A	1.392 (0.048)	1.452 (0.021)	0.000 (0.000)	-0.112 (0.028)	0.112 (0.039)			
2K040805	SE	4A	0.664 (0.020)	1.354 (0.031)	0.000 (0.000)	-1.071 (0.061)	1.071 (0.071)			
2K040806	SE	5A	0.906 (0.051)	1.580 (0.056)	0.000 (0.000)					
2K040901	SG	1A	0.859 (0.027)	-0.684 (0.028)	0.000 (0.000)					
2K040902	SG	2A	0.507 (0.017)	1.148 (0.033)	0.000 (0.000)	0.880 (0.037)	-0.880 (0.057)			
2K040903	SG	3A	0.596 (0.018)	0.540 (0.023)	0.000 (0.000)	0.649 (0.033)	-0.649 (0.040)			
2K040904	SG	4A	0.712 (0.026)	0.885 (0.034)	0.000 (0.000)					
2K040905	SG	5A	0.493 (0.020)	-0.262 (0.037)	0.000 (0.000)					
2K041001	SG	6	0.576 (0.038)	-0.346 (0.114)	0.194 (0.036)					
2K041002	SG	7A	0.557 (0.015)	-0.115 (0.021)	0.000 (0.000)	-0.166 (0.041)	0.166 (0.039)			
2K041003	SG	8	1.227 (0.062)	-1.026 (0.060)	0.251 (0.031)					
2K041004	SG	9A	0.560 (0.016)	0.648 (0.023)	0.000 (0.000)	-0.008 (0.037)	0.008 (0.044)			
2K041101	SG	10A	0.637 (0.027)	1.221 (0.047)	0.000 (0.000)					
2K041201	SG	11A	0.521 (0.013)	1.006 (0.028)	0.000 (0.000)	-0.750 (0.047)	0.750 (0.056)			
2K041202	SG	12A	0.628 (0.020)	1.442 (0.035)	0.000 (0.000)	-0.377 (0.042)	0.377 (0.058)			
2K041501	SJ	1	0.266 (0.027)	-2.552 (0.377)	0.254 (0.054)					
2K041601	SJ	2	0.811 (0.092)	0.683 (0.094)	0.433 (0.026)					
2K041701	SJ	3	1.428 (0.140)	2.180 (0.099)	0.196 (0.007)					
2K041801	SJ	4	0.845 (0.060)	-0.393 (0.099)	0.316 (0.036)					
2K041802	SJ	5	0.736 (0.067)	1.000 (0.066)	0.192 (0.021)					
2K042201	SJ	10A	0.766 (0.049)	2.794 (0.116)	0.000 (0.000)	0.269 (0.052)	-0.269 (0.172)			
2K042701	SK	1	0.816 (0.063)	0.181 (0.080)	0.269 (0.029)					
2K042702	SK	2	1.086 (0.066)	0.417 (0.042)	0.180 (0.018)					
2K042703	SK	3	0.821 (0.065)	0.405 (0.073)	0.255 (0.026)					
2K043101	SK	7A	0.587 (0.016)	0.071 (0.022)	0.000 (0.000)	-0.285 (0.045)	0.285 (0.045)			
2K043102	SK	8A	0.371 (0.011)	0.122 (0.025)	0.000 (0.000)	-0.473 (0.080)	1.103 (0.079)	-0.630 (0.069)		
2K043103	SK	9A	0.550 (0.028)	2.329 (0.089)	0.000 (0.000)	0.120 (0.052)	-0.120 (0.109)			
2K043301	SK	11	0.865 (0.069)	0.700 (0.058)	0.218 (0.021)					
2K047001	SL	3	0.395 (0.056)	1.134 (0.185)	0.246 (0.042)					
2K047301	SL	6A	0.617 (0.022)	1.118 (0.032)	0.000 (0.000)	0.378 (0.037)	-0.378 (0.053)			
2K047601	SL	9	0.844 (0.105)	2.104 (0.114)	0.087 (0.010)					
2K048301	SM	2	0.646 (0.047)	-0.294 (0.114)	0.229 (0.037)					
2K048701	SM	6	1.048 (0.108)	0.873 (0.061)	0.410 (0.018)					
2K048801	SM	7	0.575 (0.078)	1.711 (0.109)	0.182 (0.023)					
2K049401	SM	13A	0.587 (0.019)	1.810 (0.036)	0.000 (0.000)	1.435 (0.036)	-1.435 (0.093)			
2K049402	SM	14A	0.433 (0.017)	0.138 (0.031)	0.000 (0.000)	0.693 (0.054)	-0.693 (0.055)			
2K0494CL	SM		0.182 (0.005)	2.056 (0.053)	0.000 (0.000)	-10.827 (0.508)	8.638 (0.534)	3.454 (0.225)	0.125 (0.217)	-1.389 (0.307)

Table D-1 (continued)

IRT Parameters for Science Items

Earth Science, Grade 8

NAEP ID	Block	Item	a_j (s.e.)	b_j (s.e.)	c_j (s.e.)	d_{j1} (s.e.)	d_{j2} (s.e.)	d_{j3} (s.e.)	d_{j4} (s.e.)	d_{j5} (s.e.)
2K036101	SN	7A	0.487 (0.011)	0.246 (0.029)	0.000 (0.000)	1.776 (0.043)	-1.776 (0.048)			
2K0364CL	SN		0.540 (0.017)	2.429 (0.035)	0.000 (0.000)	1.597 (0.036)	0.651 (0.055)	-2.248 (0.242)		
2K036402	SN	11A	0.359 (0.016)	1.927 (0.073)	0.000 (0.000)	0.265 (0.054)	-0.265 (0.087)			
2K037401	SO	5A	0.683 (0.029)	1.929 (0.055)	0.000 (0.000)	0.277 (0.039)	-0.277 (0.080)			
2K037601	SO	7A	0.676 (0.022)	0.469 (0.022)	0.000 (0.000)	0.451 (0.034)	-0.451 (0.039)			
2K037801	SO	10	0.885 (0.072)	0.803 (0.055)	0.221 (0.019)					
2K038201	SO	15A	0.386 (0.023)	3.638 (0.135)	0.000 (0.000)	1.409 (0.062)	-1.409 (0.233)			
2K043801	ST	2	0.999 (0.075)	0.215 (0.065)	0.329 (0.024)					
2K044001	ST	4	0.724 (0.217)	3.442 (0.572)	0.120 (0.010)					
2K044101	ST	5A	1.060 (0.040)	0.892 (0.028)	0.000 (0.000)					
2K044401	ST	8A	0.590 (0.027)	-1.513 (0.066)	0.000 (0.000)					
2K044501	ST	9	0.683 (0.090)	1.715 (0.098)	0.190 (0.019)					
2K044801	ST	12	0.275 (0.030)	0.279 (0.222)	0.264 (0.036)					
2K045701	SU	6A	0.476 (0.018)	0.110 (0.029)	0.000 (0.000)	0.627 (0.050)	-0.627 (0.050)			
2K046101	SU	10	0.963 (0.076)	0.405 (0.063)	0.314 (0.023)					

Table D-2
IRT Parameters for Science Items
Physical Science, Grade 8

NAEP ID	Block	Item	a_j (s.e.)	b_j (s.e.)	c_j (s.e.)	d_{j1} (s.e.)	d_{j2} (s.e.)	d_{j3} (s.e.)	d_{j4} (s.e.)	d_{j5} (s.e.)
1K040601	SC	1A	0.662 (0.021)	-0.205 (0.025)	0.000 (0.000)					
1K0406CL	SC		0.482 (0.011)	-0.684 (0.015)	0.000 (0.000)	1.919 (0.094)	0.681 (0.059)	-0.203 (0.047)	-1.066 (0.046)	-1.330 (0.047)
1K040603	SC	3A	0.356 (0.007)	-0.549 (0.029)	0.000 (0.000)	-2.751 (0.089)	2.751 (0.085)			
1K040604	SC	4A	0.478 (0.011)	0.381 (0.021)	0.000 (0.000)	-1.104 (0.050)	1.104 (0.052)			
1K040605	SC	5A	0.499 (0.016)	0.457 (0.024)	0.000 (0.000)	0.436 (0.039)	-0.436 (0.043)			
1K040606	SC	6A	0.584 (0.012)	-0.295 (0.020)	0.000 (0.000)	2.436 (0.065)	-1.439 (0.039)	-0.996 (0.046)		
1K040702	SD	3A	0.577 (0.016)	-0.623 (0.023)	0.000 (0.000)	0.591 (0.040)	-0.591 (0.031)			
1K040704	SD	5A	0.773 (0.019)	-0.696 (0.019)	0.000 (0.000)	0.169 (0.034)	-0.169 (0.026)			
1K040705	SD	6A	0.229 (0.017)	2.002 (0.159)	0.000 (0.000)					
1K031309	SF	6G	0.216 (0.005)	-3.396 (0.100)	0.000 (0.000)	-6.808 (0.441)	2.558 (0.429)	4.250 (0.153)		
1K031302	SF	7A	0.574 (0.021)	2.158 (0.051)	0.000 (0.000)	0.801 (0.033)	-0.801 (0.085)			
1K031305	SF	8A	0.612 (0.026)	2.664 (0.069)	0.000 (0.000)	0.826 (0.036)	-0.826 (0.124)			
1K031306	SF	9A	0.720 (0.043)	2.557 (0.116)	0.000 (0.000)					
1K041901	SJ	6A	0.463 (0.019)	2.638 (0.060)	0.000 (0.000)	1.630 (0.046)	-1.630 (0.150)			
1K042301	SJ	11	0.408 (0.084)	1.944 (0.206)	0.316 (0.037)					
1K042401	SJ	12	0.611 (0.153)	2.855 (0.352)	0.241 (0.017)					
1K042501	SJ	13	1.131 (0.126)	2.009 (0.096)	0.217 (0.009)					
1K042801	SK	4	1.212 (0.104)	1.276 (0.043)	0.178 (0.012)					
1K042901	SK	5	1.666 (0.131)	1.918 (0.068)	0.240 (0.008)					
1K043201	SK	10	0.405 (0.041)	0.090 (0.195)	0.251 (0.044)					
1K043501	SK	13A	0.642 (0.017)	-0.015 (0.026)	0.000 (0.000)	1.324 (0.041)	-1.324 (0.039)			
1K043601	SK	14A	0.624 (0.023)	1.247 (0.033)	0.000 (0.000)	0.864 (0.036)	-0.864 (0.061)			
1K043602	SK	15A	0.496 (0.029)	2.987 (0.160)	0.000 (0.000)	-0.954 (0.097)	0.954 (0.179)			
1K043603	SK	16A	0.639 (0.029)	-1.080 (0.055)	0.000 (0.000)					
1K046801	SL	1	0.480 (0.053)	0.470 (0.152)	0.239 (0.040)					
1K046901	SL	2	1.024 (0.065)	0.308 (0.047)	0.199 (0.020)					
1K047101	SL	4	1.267 (0.076)	0.286 (0.039)	0.228 (0.018)					
1K047201	SL	5A	0.536 (0.014)	0.411 (0.019)	0.000 (0.000)	0.257 (0.045)	-1.083 (0.067)	0.826 (0.069)		
1K047401	SL	7A	0.556 (0.017)	0.971 (0.030)	0.000 (0.000)	-0.557 (0.048)	0.557 (0.058)			
1K047901	SL	12A	0.685 (0.023)	0.843 (0.025)	0.000 (0.000)	0.300 (0.034)	-0.300 (0.044)			
1K048201	SM	1	0.945 (0.069)	0.794 (0.046)	0.172 (0.017)					
1K048501	SM	4	0.901 (0.095)	1.169 (0.063)	0.302 (0.018)					
1K048601	SM	5A	0.429 (0.016)	1.059 (0.040)	0.000 (0.000)	1.052 (0.049)	-1.052 (0.070)			
1K049001	SM	9A	0.688 (0.047)	2.945 (0.149)	0.000 (0.000)	-2.435 (0.241)	2.435 (0.290)			
1K049101	SM	10	0.542 (0.037)	-0.434 (0.119)	0.172 (0.036)					
1K035401	SN	1	1.135 (0.129)	1.292 (0.059)	0.451 (0.013)					
1K035601	SN	3A	0.289 (0.012)	2.014 (0.071)	0.000 (0.000)	1.244 (0.060)	-1.244 (0.104)			
1K035701	SN	4	0.843 (0.050)	0.440 (0.048)	0.165 (0.019)					
1K035801	SN	5A	0.494 (0.013)	-0.676 (0.027)	0.000 (0.000)	-0.343 (0.051)	0.343 (0.043)			
1K036201	SN	8	0.985 (0.052)	0.457 (0.036)	0.135 (0.015)					
1K036301	SN	9A	0.489 (0.010)	0.475 (0.021)	0.000 (0.000)	-1.532 (0.058)	1.532 (0.060)			
1K036601	SN	13	0.535 (0.036)	-0.348 (0.122)	0.188 (0.037)					
1K036901	SN	16	0.795 (0.042)	-0.455 (0.071)	0.179 (0.028)					
1K037301	SO	4A	0.444 (0.012)	0.945 (0.032)	0.000 (0.000)	-1.738 (0.075)	1.738 (0.082)			
1K037501	SO	6A	0.737 (0.032)	2.096 (0.063)	0.000 (0.000)	-0.155 (0.047)	0.155 (0.088)			
1K037701	SO	8A	0.320 (0.014)	-3.141 (0.087)	0.000 (0.000)	0.440 (0.329)	1.512 (0.162)	-1.951 (0.064)		
1K037703	SO	9A	0.482 (0.018)	0.792 (0.033)	0.000 (0.000)	0.478 (0.045)	-0.478 (0.057)			
1K037901	SO	11	0.642 (0.106)	2.317 (0.175)	0.162 (0.016)					

Table D-2 (continued)

IRT Parameters for Science Items

Physical Science, Grade 8

NAEP ID	Block	Item	a_j (s.e.)	b_j (s.e.)	c_j (s.e.)	d_{j1} (s.e.)	d_{j2} (s.e.)	d_{j3} (s.e.)	d_{j4} (s.e.)	d_{j5} (s.e.)
1K038301	SO	16A	0.782 (0.020)	1.393 (0.019)	0.000 (0.000)	2.378 (0.039)	1.021 (0.032)	-0.957 (0.061)	-2.442 (0.287)	
1K043901	ST	3	0.651 (0.062)	0.254 (0.112)	0.301 (0.034)					
1K044201	ST	6A	0.366 (0.011)	-0.453 (0.042)	0.000 (0.000)	1.856 (0.074)	-1.856 (0.062)			
1K044301	ST	7A	0.384 (0.016)	-0.269 (0.035)	0.000 (0.000)	0.458 (0.064)	-0.458 (0.059)			
1K044701	ST	11	0.556 (0.065)	1.434 (0.095)	0.152 (0.024)					
1K045101	ST	15A	0.632 (0.019)	1.248 (0.027)	0.000 (0.000)	-0.183 (0.047)	0.633 (0.059)	-0.450 (0.075)		
1K045102	ST	16A	0.478 (0.019)	1.492 (0.044)	0.000 (0.000)	0.791 (0.049)	-0.732 (0.084)	-0.058 (0.122)		
1K045401	SU	3	0.967 (0.087)	0.080 (0.089)	0.472 (0.027)					
1K045501	SU	4	0.530 (0.065)	1.110 (0.113)	0.235 (0.031)					
1K045801	SU	7A	0.548 (0.016)	1.231 (0.033)	0.000 (0.000)	1.444 (0.038)	-1.444 (0.070)			
1K045901	SU	8	1.203 (0.146)	1.173 (0.061)	0.455 (0.015)					
1K046201	SU	11	0.435 (0.099)	2.430 (0.262)	0.296 (0.030)					
1K046501	SU	14A	0.706 (0.026)	1.169 (0.030)	0.000 (0.000)	0.578 (0.032)	-0.578 (0.052)			
1K046601	SU	15A	0.572 (0.021)	0.489 (0.026)	0.000 (0.000)	0.279 (0.042)	-0.279 (0.047)			

Table D-3

IRT Parameters for Science Items

Life Science, Grade 8

NAEP ID	Block	Item	a_j (s.e.)	b_j (s.e.)	c_j (s.e.)	d_{j1} (s.e.)	d_{j2} (s.e.)	d_{j3} (s.e.)	d_{j4} (s.e.)	d_{j5} (s.e.)
3K031307	SF	10A	0.851 (0.027)	1.797 (0.029)	0.000 (0.000)	0.848 (0.024)	-0.848 (0.065)			
3K031308	SF	11A	0.510 (0.026)	2.588 (0.106)	0.000 (0.000)	-0.324 (0.059)	0.324 (0.116)			
3K041301	SH	1	0.670 (0.039)	-2.634 (0.152)	0.203 (0.055)					
3K041302	SH	2	0.861 (0.046)	-1.610 (0.111)	0.244 (0.050)					
3K041303	SH	3	0.753 (0.034)	-1.268 (0.087)	0.152 (0.036)					
3K041304	SH	4	0.600 (0.037)	0.186 (0.077)	0.129 (0.025)					
3K041305	SH	5	1.212 (0.052)	-0.729 (0.046)	0.160 (0.024)					
3K041306	SH	6A	0.642 (0.022)	-0.805 (0.036)	0.000 (0.000)					
3K041307	SH	7A	0.375 (0.017)	1.985 (0.076)	0.000 (0.000)	0.153 (0.053)	-0.153 (0.087)			
3K041401	SH	8A	0.896 (0.022)	0.758 (0.018)	0.000 (0.000)	0.931 (0.023)	-0.931 (0.032)			
3K041402	SH	9A	0.942 (0.024)	1.665 (0.021)	0.000 (0.000)	1.252 (0.021)	-1.252 (0.069)			
3K041403	SH	10A	0.276 (0.014)	4.278 (0.134)	0.000 (0.000)	2.087 (0.068)	-2.087 (0.233)			
3K031601	SI	1	0.872 (0.048)	-1.707 (0.118)	0.256 (0.055)					
3K031602	SI	2A	0.386 (0.014)	-0.582 (0.033)	0.000 (0.000)	0.662 (0.058)	-0.662 (0.049)			
3K031603	SI	3A	0.616 (0.025)	-1.702 (0.060)	0.000 (0.000)					
3K031604	SI	4A	0.434 (0.017)	1.923 (0.056)	0.000 (0.000)	0.847 (0.043)	-0.847 (0.084)			
3K031605	SI	5	0.351 (0.025)	-1.605 (0.216)	0.225 (0.045)					
3K031606	SI	6A	0.421 (0.012)	1.161 (0.035)	0.000 (0.000)	1.536 (0.043)	-1.536 (0.068)			
3K031610	SI	7A	0.617 (0.014)	0.222 (0.024)	0.000 (0.000)	1.427 (0.035)	-1.427 (0.038)			
3K031607	SI	8A	0.459 (0.012)	-0.761 (0.024)	0.000 (0.000)	0.169 (0.075)	1.100 (0.059)	-1.269 (0.042)		
3K031608	SI	9A	0.431 (0.011)	1.994 (0.037)	0.000 (0.000)	2.415 (0.041)	-2.415 (0.109)			
3K031609	SI	10A	0.341 (0.012)	1.267 (0.049)	0.000 (0.000)	-0.446 (0.060)	0.446 (0.076)			
3K031611	SI	11A	0.344 (0.023)	2.371 (0.148)	0.000 (0.000)					
3K031612	SI	12	0.845 (0.056)	-0.346 (0.092)	0.333 (0.032)					
3K031613	SI	13A	0.601 (0.024)	2.360 (0.075)	0.000 (0.000)	-1.053 (0.070)	1.053 (0.106)			
3K042001	SJ	7A	0.539 (0.022)	1.751 (0.049)	0.000 (0.000)	0.819 (0.041)	-0.819 (0.082)			
3K042101	SJ	8A	0.229 (0.020)	8.626 (0.390)	0.000 (0.000)	4.265 (0.112)	-4.265 (1.149)			
3K042102	SJ	9A	0.721 (0.021)	1.220 (0.027)	0.000 (0.000)	1.183 (0.030)	-1.183 (0.059)			
3K042601	SJ	14A	0.494 (0.013)	0.274 (0.025)	0.000 (0.000)	-0.783 (0.056)	0.783 (0.058)			
3K042602	SJ	15A	0.608 (0.022)	1.094 (0.032)	0.000 (0.000)	0.690 (0.037)	-0.690 (0.056)			
3K042603	SJ	16A	0.549 (0.032)	1.717 (0.085)	0.000 (0.000)					
3K043001	SK	6A	0.268 (0.019)	3.652 (0.209)	0.000 (0.000)	0.928 (0.085)	-0.928 (0.202)			
3K043401	SK	12	0.277 (0.027)	-0.588 (0.236)	0.217 (0.040)					
3K047501	SL	8	0.209 (0.036)	3.651 (0.544)	0.217 (0.024)					
3K047701	SL	10	0.577 (0.143)	3.170 (0.440)	0.120 (0.014)					
3K047801	SL	11	1.026 (0.095)	1.266 (0.048)	0.180 (0.013)					
3K048001	SL	13A	0.838 (0.019)	0.921 (0.026)	0.000 (0.000)	1.818 (0.030)	-1.818 (0.062)			
3K048101	SL	14A	0.368 (0.016)	1.975 (0.073)	0.000 (0.000)	0.035 (0.066)	-1.184 (0.139)	1.149 (0.179)		
3K048102	SL	15A	0.716 (0.059)	3.799 (0.176)	0.000 (0.000)	1.035 (0.061)	-1.035 (0.556)			
3K048103	SL	16A	0.494 (0.026)	0.801 (0.054)	0.000 (0.000)					
3K048401	SM	3	0.567 (0.059)	0.725 (0.105)	0.205 (0.032)					
3K048901	SM	8A	0.461 (0.017)	0.546 (0.032)	0.000 (0.000)	0.621 (0.048)	-0.621 (0.056)			
3K049201	SM	11	0.177 (0.023)	0.297 (0.297)	0.220 (0.034)					
3K049301	SM	12A	0.296 (0.034)	4.685 (0.512)	0.000 (0.000)					
3K035501	SN	2	0.506 (0.029)	-1.910 (0.174)	0.204 (0.054)					
3K035901	SN	6A	0.490 (0.017)	1.554 (0.045)	0.000 (0.000)	0.081 (0.042)	-0.081 (0.064)			
3K036501	SN	12	0.485 (0.066)	1.094 (0.134)	0.289 (0.035)					
3K036701	SN	14A	0.412 (0.020)	2.729 (0.116)	0.000 (0.000)	-0.496 (0.065)	0.496 (0.120)			

Table D-3 (continued)

IRT Parameters for Science Items

Life Science, Grade 8

NAEP ID	Block	Item	a_j (s.e.)	b_j (s.e.)	c_j (s.e.)	d_{j1} (s.e.)	d_{j2} (s.e.)	d_{j3} (s.e.)	d_{j4} (s.e.)	d_{j5} (s.e.)
3K036801	SN	15A	0.595 (0.018)	1.436 (0.035)	0.000 (0.000)	-0.254 (0.039)	0.254 (0.055)			
3K037101	SO	2	0.256 (0.026)	-3.213 (0.430)	0.220 (0.059)					
3K037201	SO	3	0.835 (0.053)	-0.886 (0.109)	0.262 (0.043)					
3K038001	SO	12	0.780 (0.105)	1.693 (0.089)	0.211 (0.017)					
3K038002	SO	13	0.965 (0.092)	1.371 (0.053)	0.166 (0.013)					
3K038101	SO	14A	0.604 (0.041)	2.504 (0.137)	0.000 (0.000)					
3K043701	ST	1	0.602 (0.041)	-0.934 (0.146)	0.213 (0.048)					
3K044601	ST	10	0.130 (0.029)	7.604 (1.606)	0.147 (0.020)					
3K044901	ST	13A	0.436 (0.011)	0.730 (0.039)	0.000 (0.000)	1.981 (0.052)	-1.981 (0.072)			
3K045001	ST	14A	0.441 (0.014)	1.119 (0.039)	0.000 (0.000)	1.512 (0.048)	-1.512 (0.076)			
3K045201	SU	1	0.663 (0.042)	-1.504 (0.150)	0.232 (0.054)					
3K045301	SU	2A	0.255 (0.014)	1.315 (0.076)	0.000 (0.000)	0.744 (0.083)	-0.744 (0.109)			
3K045601	SU	5A	0.384 (0.012)	1.639 (0.040)	0.000 (0.000)	-0.504 (0.073)	1.654 (0.086)	-1.150 (0.121)		
3K046001	SU	9	0.402 (0.036)	0.447 (0.141)	0.137 (0.035)					
3K046301	SU	12A	0.346 (0.009)	1.549 (0.053)	0.000 (0.000)	3.100 (0.063)	-3.100 (0.123)			
3K046401	SU	13A	0.572 (0.029)	2.191 (0.080)	0.000 (0.000)	0.201 (0.048)	-0.201 (0.101)			
3K046701	SU	16A	0.567 (0.014)	1.157 (0.030)	0.000 (0.000)	2.136 (0.043)	-1.439 (0.067)	-0.697 (0.113)		

Appendix E

STATE ASSESSMENT PROGRAM REPORTING SUBGROUPS; COMPOSITE AND DERIVED COMMON BACKGROUND VARIABLES; AND COMPOSITE AND DERIVED REPORTING VARIABLES

REPORTING SUBGROUPS FOR THE 1996 STATE ASSESSMENT

Results for the 1996 State Assessment were reported for student subgroups defined by gender, race/ethnicity, type of location, parents' level of education, participation in the National School Lunch Program, and eligibility of Title I funding. The following explains how each of these subgroups was derived.

DSEX (Gender)

The variable SEX is the gender of the student being assessed, as taken from school records. For a few students, data for this variable was missing and was imputed by ETS after the assessment. The resulting variable DSEX contains a value for every student and is used for gender comparisons among students.

DRACE (Race/ethnicity)

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

<u>Demographic Item Number 2</u>:

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

- ⬭ I am not Hispanic
- ⬭ Mexican, Mexican American, or Chicano
- ⬭ Puerto Rican
- ⬭ Cuban
- ⬭ Other Spanish or Hispanic background

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

Demographic Item Number 1:

1. Which best describes you?

- ⬭ White (not Hispanic)
- ⬭ Black (not Hispanic)
- ⬭ Hispanic (“Hispanic” means someone who is Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or from some other Spanish or Hispanic background.)
- ⬭ Asian or Pacific Islander (“Asian or Pacific Islander” means someone who is Chinese, Japanese, Korean, Filipino, Vietnamese, or from some other Asian or Pacific Island background.)
- ⬭ American Indian or Alaskan Native (“American Indian or Alaskan Native” means someone who is from one of the American Indian tribes, or one of the original people of Alaska.)
- ⬭ Other (What?) __________

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

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

An exception in this definition of race/ethnicity was made for Hawaii. Students from Hawaii who specified Asian or Pacific Islander for Demographic Item Number 1 were categorized in the Asian or Pacific Islander race/ethnicity classification, no matter what response they gave for Demographic Item Number 2.

TOL8 (Type of Location)
TOL5
TOL3

The variable TOL8 is used by NAEP to provide information about the type of location in which schools are located. The variable is defined using population size information from the 1990 Census and the definitions of Metropolitan Statistical Areas (MSAs) as of June 1995. The variables TOL8, TOL5 and TOL3 apply only to the eighth-grade sample. There are eight categories for TOL8:

1	Large Central City	a central city of an MSA with a population greater than or equal to 400,000, or a population density greater than or equal to 6,000 persons per square mile
2	Midsize Central City	a central city of an MSA but not designated as a large city
3	Urban Fringe of Large City	a place within an MSA of a large central city and defined as urban by the U.S. Bureau of Census
4	Urban Fringe of a Midsize City	a place within an MSA of a midsize central city and defined as urban by the U.S. Bureau of Census
5	Large Town	a place not within an MSA, but with a population greater than or equal to 25,000 and defined as urban by the U.S. Bureau of Census
6	Small Town	a place not within an MSA, with a population less than 25,000, but greater than or equal to 2,500 and defined as urban by the U.S. Bureau of Census
7	Rural MSA	a place within an MSA with a population of less than 2,500 and defined as rural by the U.S. Bureau of the Census
8	Rural NonMSA	a place not within an MSA with a population of less than 2,500 and defined as rural by the U.S. Bureau of the Census

The variable TOL5 collapses the information provided in the variable TOL8 to five levels:

1 Large Central City
2 Midsize Central City
3 Urban Fringe of Large City, Urban Fringe of Midsize City, and Large Town
4 Small Town
5 Rural MSA and Rural NonMSA

The variable TOL3 is used extensively in the NAEP reports. TOL3 collapses TOL8 to three levels:

1	Central City	(Large Central City and Midsize Central City) This category includes central cities of all MSAs. Central City is a geographic term and is not synonymous with "inner city."
2	Urban Fringe/Large Town	(Urban Fringe of Large City, Urban Fringe of Midsize City, and Large Town) An Urban Fringe includes all densely settled places and areas within MSAs that are classified as urban by the Bureau of the Census. A Large Town is defined as a place outside MSAs with a population greater than or equal to 25,000.

3	Rural/Small Town	(Small Town, Rural MSA, and Rural NonMSA) Rural includes all places and areas with a population of less than 2,500 that are classified as rural by the Bureau of the Census. A Small Town is defined as a place outside MSAs with a population of less than 25,000 but greater than or equal to 2,500.

PARED (Parents' education level)

The variable PARED is derived from responses to two questions, B003501 and B003601, in the student demographic questionnaire. Students were asked to indicate the extent of their mother's education (B003501—How far in school did your mother go?) by choosing one of the following:

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

Students were asked to provide the same information about the extent of their father's education (B003601—How far in school did your father go?) by choosing one of the following:

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

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

REGION (Region of the country)

For the eighth-grade sample, results for each jurisdiction were compared to the appropriate regional results from the national component of the assessment. Jurisdictions were grouped into four geographical regions—Northeast, Southeast, Central, and West—as shown in Table E-1. All 50 states and the District of Columbia are listed, with the participants in the State Assessment shown in italic type. Territories were not assigned to a region. The part of Virginia that is included in the Washington, DC, metropolitan statistical area is included in the Northeast region; the remainder of the state is included in the Southeast region. This variable does not apply to the fourth-grade sample.

Table E-1
NAEP Geographic Regions

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

IEP (Individualized Education Plan)[1]

The variable IEP comes from the student booklet cover. A value of 1 indicates that a student has an individualized education plan while a value of 2 indicates no individualized education program.

[1] A student identified on the Administration Schedule as a student with a disability (SD) or an equivalent classification may be excluded from the assessment if: 1) the student is mainstreamed less than 50% of the time in academic subjects and is judged incapable of participating meaningfully in the assessment, or 2) the Individualized Education Plan (IEP) team or equivalent group has determined that the student is incapable of participating meaningfully in the assessment. SD/LEP students meeting the above criteria should be assessed if, in the judgment of school staff, they are capable of taking the assessment.

LEP (Limited English Proficiency)

The variable LEP comes from the student booklet cover. A value of 1 indicates that a student is considered to have limited English proficiency while a value of 2 indicates that the student does not have limited English proficiency.

TITLE1

The variable TITLE1 comes from the student booklet cover. A value of 1 indicates that a student is eligible for Title 1 funding and a value of 2 indicates that the student is not eligible for Title 1 funding.

SLUNCH
SLUNCH1

The variable SLUNCH is provided by Westat, Inc., and is used to determine if a student participates in the National School Lunch Program. The values for this variable are as follow:

1 not eligible
2 eligible for reduced price lunch
3 eligible for free lunch
4 no information available
5 school refused to provide information

The variable SLUNCH1 collapses the information provided in the variable SLUNCH to three levels:

1 eligible for free or reduced price lunch
2 not eligible for free or reduced price lunch
3 no information available

SCHTYPE

The variable SCHTYPE is provided by Westat, Inc., and is used to determine the type of school that a student attended. The values for this variable are as follow:

1 Public
2 Private
3 Catholic
4 Bureau of Indian Affairs (BIA)
5 Department of Defense Education Activity (DoDEA) schools

VARIABLES DERIVED FROM THE STUDENT, TEACHER, AND SCHOOL QUESTIONNAIRES

Several variables were formed from the systematic combination of response values for one or more items from either the student demographic questionnaire, the student science background questionnaire, the teacher questionnaire, or the school questionnaire.

HOMEEN3 (Home environment—Articles [of 4] in the home)

The variable HOMEEN3 was created from the responses to student demographic items B000901 (Does your family get a newspaper regularly?), B000903 (Is there an encyclopedia in your home?), B008801 (How many books are in your home?), collapsed to indicate whether or not there are more than 25 books in the home), and B000905 (Does your family get any magazines regularly?). The values for this variable were derived as follows:

1	0-2 types	The student responded to at least two items and answered Yes to two or fewer.
2	3 types	The student answered Yes to three items.
3	4 types	The student answered Yes to four items.
8	Omitted	The student answered fewer than two items.

VARIABLES DERIVED FROM SCIENCE ITEMS

SCITAKE

The variable SCITAKE is available for the grade 8 sample. This variable was created from items K812001 through K812006 in order to provide maximum utility of this data. For some analyses this variable was used to interpret these items as a single response item. When a student responded to more than one prompt SCITAKE was coded as a multiple response, however when the student responded to only one prompt the recoding was as follows:

1 I am not taking science this year
2 Life science
3 Physical science
4 Earth science
5 General science
6 Integrated science

BKSCOR (Booklet-Level score)

The booklet-level score is a student-level score based on the sum of the number correct for dichotomous items plus the sum of the scores on the polytomous items, where the score for a polytomous item starts from 0 for the unacceptable category. Thus, for a 4-point extended constructed-response item, scores of “no response”, “off-task”, and “unsatisfactory” are assigned an item score of 0. Scores of “partial”, “essential”, and “extensive” are assigned item scores of 1, 2, and 3, respectively. The score is computed based on all cognitive items in an individual’s assessment booklet. This is available for the eighth-grade sample.

LOGIT (Logit percent correct within booklet)

In order to compute the LOGIT score, a percent correct within booklet was first computed. This score was based on the ratio of the booklet score (BKSCOR) over the maximum booklet score. The percent correct score was set to .0001 if no items were answered correctly; if BKSCOR equaled the maximum booklet score, the percent correct score was set to .9999.
A logit score, LOGIT, was calculated for each student by the following formula:

$$LOGIT = ln\left[\frac{PCTCOR}{1 - PCTCOR}\right]$$

LOGIT was then restricted to a value x, such that $-3 \leq x \leq 3$. After computing LOGIT for each student, the mean and standard deviation was calculated for each booklet using the reporting sample as the first step in standardizing the logit scores. The standardized logit score, ZLOGIT, was then calculated for each student by the following formula:

$$ZSCORE = \left[\frac{LOGIT - mean}{standard\ deviation}\right]$$

This is available for the eighth-grade sample.

NORMIT (Normit Gaussian score)
SCHNORM (School-level mean Gaussian score)

The normit score is a student-level Gaussian score based on the inverse normal transformation of the mid-percentile rank of a student’s number-correct booklet score within that booklet. The normit scores were used to decide collapsing of variables, finalize conditioning coding, and check the results of scaling.

The number-correct is based on the number of dichotomous items answered correctly plus the score obtained on extended constructed-response items. The distributions of normit scores were constructed from the reporting sample using the overall reporting weight. The minimum and maximum normits were set to -3.5 and +3.5, respectively. The mid-percentile rank is based on the formula:

$$\frac{CF(i) + CF(i - 1)}{2N}$$

where CF(i) is the cumulative frequency at i items correct and N is the total sample size. If i = 0 then

$$\frac{CF(0) + \frac{CF(1)}{2}}{2N}$$

A school-level normit, SCHNORM was also created; this was the mean normit across all main assessment science booklets administered in a school. These school-level mean normit scores were used in conditioning procedures to take into account differences in school proficiency. For each school, the weighted mean (the within school modular student weight, corrected for student nonresponse) of the normits for the students in that school was calculated. Each student was then assigned that mean as his or her school-level mean normit score value.

VARIABLES RELATED TO PROFICIENCY SCALING

Proficiency Score Variables

Item response theory (IRT) was used to estimate average science proficiency for each jurisdiction and for various subpopulations, based on students' performance on the set of science items they received. IRT provides a common scale on which performance can be reported for the nation, jurisdiction, and subpopulations, even when all students do not answer the same set of questions. This common scale makes it possible to report on relationships between students' characteristics (based on their responses to the background questions) and their overall performance in the assessment.

A scale ranging from 0 to 300 was created to report performance for each of the three fields of science: *earth science, physical science, and life science*. Each scale was based on the distribution of student performance across all three grades assessed in the 1996 national assessment (grades 4, 8, and 12) and had a mean of 150 and a standard deviation of 50. A composite scale was created as an overall measure of students' science proficiency. The composite scale was a weighted average of the three fields of science scales, where the weight for each field was proportional to the relative importance assigned to the field as specified in the science objectives. Although the items comprising each scale were identical to those used for the national program, the item parameters for the State Assessment scales were estimated from the combined data from all jurisdictions participating in the State Assessment.

Scale proficiency estimates were obtained for all students assessed in the State Assessment. The NAEP methods use random draws ("plausible values") from estimated proficiency distributions to compute population statistics. Plausible values are not optimal estimates of individual proficiency; instead, they serve as intermediate values to be used in estimating population characteristics. Chapter 8 provides further details on the computation and use of plausible values.

The proficiency score (plausible value) variables are provided on the student data files for each of the scales and are named as shown in Table E-2.

Table E-2
Scaling Variables for the 1996 State Assessment Samples

Fields of Science Scale	Data Variables
Earth Science	SRPS21 to SRPS25
Physical Science	SRPS11 to SRPS15
Life Science	SRPS31 to SRPS35
Composite	SRPCM1 to SRPCM5

QUALITY EDUCATION DATA VARIABLES (QED)

The data files contain several variables obtained from information supplied by Quality Education Data, Inc. (QED). QED maintains and updates annually lists of schools showing grade span, total enrollment, instructional dollars per pupil, and other information for each school. These data variables are retained on both the school and student files and are identified in the data layouts by ("QED") in the SHORT LABEL field.

Most of the QED variables are defined sufficiently in the data codebooks. Explanations of others are provided below.

ORSHPT is the Orshansky Percentile, an indicator of relative wealth that specifies the percentage of school-age children in a district who fall below the poverty line.

IDP represents, at the school district level, dollars per student spent for textbooks and supplemental materials. The range code for instructional dollars spent per pupil excluding teacher salaries are:

0 = Unclassified
1 = Under $10
2 = $10-49
3 = $50-99
4 = $100-149
5 = $150-299
6 = $300-399
7 = $400-499
8 = $500-999
9 = $1000 +

ADULTED indicates whether or not adult education courses are offered at the school site.

URBAN defines the school's urbanization: urban (central city); suburban (area surrounding central city, but still located within the counties constituting the metropolitan statistical area); or rural (area outside any metropolitan statistical area).

Appendix F

SAMPLE DESIGN AND SELECTION TABLES

John Burke and James Green
Westat, Inc.

This appendix contains the urbanization classifications used within each jurisdiction for grade 8 (Table F-1). Urbanization classification was created by collapsing type of location categories as necessary and according to specific rules until each urbanization stratus included a minimum of 10 percent of eligible students in the participating jurisdiction.

Also included in this appendix is information on metro area status stratification for the participating jurisdictions (Table F-2). All schools in the sampling frame were assigned a metro area status based on their Federal Information Processing Standards (FIPS) county code and Office of Management and Budget (OMB) Metropolitan Area Definitions as of June 30, 1993. This field indicated if a school was located within a metropolitan area or not.

Tables F-3 and F-4 include information about the number of substitutes provided in each jurisdiction. Of the 47 participating jurisdictions, 41 were provided with at least one substitute at grade 8. Among jurisdictions receiving no substitutes, the majority had 100 percent participation from the original sample.

Tables F-5 through F-7 show the number of schools in the fourth- and eighth-grade science samples, together with school response rates observed within participating jurisdictions. The tables also show the number of substitutes in each jurisdiction that were associated with a nonparticipating original school selection, and the number of those that participated.

This appendix also contains the distribution of the student samples and response rates by grade, school type, and jurisdiction in Tables F-8 through F-10.

Table F-1
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Alaska				
Large	Large	Mid-size Central City/Urban Fringe	None	8
Small	Small	Small Town	None	2
Small	Small	Rural	None	79
Small	Large	Small Town	None	16
Small	Large	Rural	None	16
Alabama				
Small	Small	Rural	High	1
Small	Large	Mid-size Central City	Low	7
Small	Large	Mid-size Central City	Medium	7
Small	Large	Mid-size Central City	High	7
Small	Large	Urban Fringe of Mid-size Central City	Low	10
Small	Large	Urban Fringe of Mid-size Central City	Medium	10
Small	Large	Urban Fringe of Mid-size Central City	High	9
Small	Large	Large/Small Town	Low	9
Small	Large	Large/Small Town	Medium	10
Small	Large	Large/Small Town	High	9
Small	Large	Rural	Low	11
Small	Large	Rural	Medium	9
Small	Large	Rural	High	11

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Arkansas				
Small	Small	Rural	Low	1
Small	Small	Rural	Medium	1
Small	Small	Rural	High	1
Small	Large	Mid-size Central City/Urban Fringe	Low	8
Small	Large	Mid-size Central City/Urban Fringe	Medium	9
Small	Large	Mid-size Central City/Urban Fringe	High	7
Small	Large	Large/Small Town	Low	17
Small	Large	Large/Small Town	Medium	16
Small	Large	Large/Small Town	High	16
Small	Large	Rural	Low	12
Small	Large	Rural	Medium	13
Small	Large	Rural	High	12
Arizona				
Small	Small	Mid-size Central City	Medium	1
Small	Small	Large/Small Town/Rural	Medium	2
Small	Large	Large Central City	Low	7
Small	Large	Large Central City	Medium	9
Small	Large	Large Central City	High	9
Small	Large	Mid-size Central City	Low	11
Small	Large	Mid-size Central City	Medium	10
Small	Large	Mid-size Central City	High	10
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	5
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	6
Small	Large	Urban Fringe of Large/Mid-size Central City	High	6
Small	Large	Large/Small Town/Rural	Low	11
Small	Large	Large/Small Town/Rural	Medium	12
Small	Large	Large/Small Town/Rural	High	10

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
California				
Small	Small	Large Central City	Medium	1
Small	Small	Large/Small Town/Rural	Low	1
Small	Large	Large Central City	Low	7
Small	Large	Large Central City	Medium	7
Small	Large	Large Central City	High	6
Small	Large	Mid-size Central City	Low	6
Small	Large	Mid-size Central City	Medium	6
Small	Large	Mid-size Central City	High	6
Small	Large	Urban Fringe of Large Central City	Low	11
Small	Large	Urban Fringe of Large Central City	Medium	11
Small	Large	Urban Fringe of Large Central City	High	11
Small	Large	Urban Fringe of Mid-size Central City	Low	4
Small	Large	Urban Fringe of Mid-size Central City	Medium	5
Small	Large	Urban Fringe of Mid-size Central City	High	4
Small	Large	Large/Small Town/Rural	Low	8
Small	Large	Large/Small Town/Rural	Medium	7
Small	Large	Large/Small Town/Rural	High	8

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Colorado				
Small	Small	Urban Fringe of Large/Mid-size Central City	Medium	1
Small	Small	Rural	Low	2
Small	Small	Rural	Medium	1
Small	Small	Rural	High	1
Small	Large	Large Central City	Low Hispanic/Low Black	2
Small	Large	Large Central City	Low Hispanic/High Black	3
Small	Large	Large Central City	High Hispanic/Low Black	3
Small	Large	Large Central City	High Hispanic/High Black	3
Small	Large	Mid-size Central City	Low	8
Small	Large	Mid-size Central City	Medium	8
Small	Large	Mid-size Central City	High	8
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	12
Small	Large	Urban Fringe of Large./Mid-size Central City	Medium	11
Small	Large	Urban Fringe of Large/Mid-size Central City	High	12
Small	Large	Large/Small Town	Low	6
Small	Large	Large/Small Town	Medium	7
Small	Large	Large/Small Town	High	7
Small	Large	Rural	Low	5
Small	Large	Rural	Medium	6
Small	Large	Rural	High	6

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Connecticut				
Small	Large	Large Central City	Low Black/Low Hispanic	2
Small	Large	Large Central City	Low Black/High Hispanic	2
Small	Large	Large Central City	High Black/Low Hispanic	3
Small	Large	Large Central City	High Black/High Hispanic	2
Small	Large	Mid-size Central City	Low	7
Small	Large	Mid-size Central City	Medium	6
Small	Large	Mid-size Central City	High	7
Small	Large	Urban Fringe of Large Central City	Low	6
Small	Large	Urban Fringe of Large Central City	Medium	6
Small	Large	Urban Fringe of Large Central City	High	6
Small	Large	Urban Fringe of Mid-size Central City	None	16
Small	Large	Large/Small Town/Rural	None	40
District of Columbia				
Large	Large	Large Central City	Low	10
Large	Large	Large Central City	Medium	12
Large	Large	Large Central City	High	11
Large	Small	Large Central City	Low	1

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Delaware				
Large	Large	Urban Fringe of Mid-size Central City	Low	1
Large	Large	Urban Fringe of Mid-size Central City	Medium	1
Large	Large	Urban Fringe of Mid-size Central City	High	1
Small	Large	Mid-size Central City	Low	1
Small	Large	Mid-size Central City	Medium	1
Small	Large	Mid-size Central City	High	1
Small	Large	Urban Fringe of Mid-size Central city	Low	1
Small	Large	Urban Fringe of Mid-size Central city	Medium	4
Small	Large	Urban Fringe of Mid-size Central city	High	2
Small	Large	Small Town	Low	2
Small	Large	Small Town	Medium	2
Small	Large	Small Town	High	2
Small	Large	Rural	Low	4
Small	Large	Rural	Medium	5
Small	Large	Rural	High	3

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Florida				
Small	Small	Large/Small Town/Rural	Low	1
Small	Large	Large Central City	Low Black/Low Hispanic	3
Small	Large	Large Central City	Low Black/High Hispanic	4
Small	Large	Large Central City	High Black/Low Hispanic	3
Small	Large	Large Central City	High Black/High Hispanic	4
Small	Large	Mid-Size Central City	Low	11
Small	Large	Mid-Size Central City	Medium	12
Small	Large	Mid-Size Central City	High	11
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	14
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	13
Small	Large	Urban Fringe of Large/Mid-size Central City	High	14
Small	Large	Large/Small Town/Rural	Low	8
Small	Large	Large/Small Town/Rural	Medium	6
Small	Large	Large/Small Town/Rural	High	6

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Georgia				
Small	Small	Large/Small Town	Low	1
Small	Large	Large/Mid-size Central City	Low	6
Small	Large	Large/Mid-size Central City	Medium	6
Small	Large	Large/Mid-size Central City	High	7
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	11
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	12
Small	Large	Urban Fringe of Large/Mid-size Central City	High	13
Small	Large	Large/Small Town	Low	12
Small	Large	Large/Small Town	Medium	11
Small	Large	Large/Small Town	High	12
Small	Large	Rural	Low	5
Small	Large	Rural	Medium	6
Small	Large	Rural	High	6
Hawaii				
Large	Large	Mid-size Central City	None	12
Large	Large	Urban Fringe of Mid-size Central City	None	20
Large	Large	Small Town/Rural	None	20
Large	Small	Small Town/Rural	None	1
Iowa				
Small	Small	Large/Small Town	None	1
Small	Small	Rural	None	2
Small	Large	Urban Fringe of Large/Mid-size Central City	None	33
Small	Large	Large/Small Town	None	39
Small	Large	Rural	None	41

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Indiana				
Small	Large	Large/Mid-size Central City	Low	9
Small	Large	Large/Mid-size Central City	Medium	9
Small	Large	Large/Mid-size Central City	High	9
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	5
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	3
Small	Large	Urban Fringe of Large/Mid-size Central City	High	5
Small	Large	Urban Fringe of Large/Mid-size Central City	None	13
Small	Large	Large/Small Town	None	33
Small	Large	Rural	None	21
Kentucky				
Small	Small	Rural	None	1
Small	Large	Mid-size Central City	Low	6
Small	Large	Mid-size Central City	Medium	6
Small	Large	Mid-size Central City	High	5
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	6
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	5
Small	Large	Urban Fringe of Large/Mid-size Central City	High	5
Small	Large	Large/Small Town	None	43
Small	Large	Rural	None	33

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Louisiana				
Small	Small	Large/Small Town	Low	1
Small	Small	Rural	High	1
Small	Large	Large/Mid-size Central City	Low	10
Small	Large	Large/Mid-size Central City	Medium	10
Small	Large	Large/Mid-size Central City	High	11
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	9
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	8
Small	Large	Urban Fringe of Large/Mid-size Central City	High	7
Small	Large	Large/Small Town	Low	12
Small	Large	Large/Small Town	Medium	11
Small	Large	Large/Small Town	High	11
Small	Large	Rural	Low	8
Small	Large	Rural	Medium	8
Small	Large	Rural	High	8
Massachusetts				
Small	Small	Small Town	None	1
Small	Large	Large/Mid-size Central City	Low	8
Small	Large	Large/Mid-size Central City	Medium	9
Small	Large	Large/Mid-size Central City	High	8
Small	Large	Urban Fringe of Large Central City	Low	7
Small	Large	Urban Fringe of Large Central City	Medium	7
Small	Large	Urban Fringe of Large Central City	High	7
Small	Large	Urban Fringe of Mid-size Central City	None	13
Small	Large	Large Town	None	18
Small	Large	Small Town	None	19
Small	Large	Rural	None	11

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Maryland				
Small	Large	Large/Mid-size Central City	Low	6
Small	Large	Large/Mid-size Central City	Medium	5
Small	Large	Large/Mid-size Central City	High	5
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	22
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	21
Small	Large	Urban Fringe of Large/Mid-size Central City	High	21
Small	Large	Large/Small Town/Rural	Low	9
Small	Large	Large/Small Town/Rural	Medium	8
Small	Large	Large/Small Town/Rural	High	9
Maine				
Small	Small	Mid-size Central City/Urban Fringe	None	1
Small	Small	Small Town	None	2
Small	Small	Rural	None	13
Small	Large	Mid-size Central City/Urban Fringe	None	16
Small	Large	Small Town	None	53
Small	Large	Rural	None	33
Michigan				
Small	Small	Rural	None	1
Small	Large	Large/Mid-size Central City	Low	7
Small	Large	Large/Mid-size Central City	Medium	8
Small	Large	Large/Mid-size Central City	High	7
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	11
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	12
Small	Large	Urban Fringe of Large/Mid-size Central City	High	12
Small	Large	Large/Small Town	None	32
Small	Large	Rural	None	17

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Minnesota				
Small	Small	Rural	None	1
Small	Large	Large/Mid-size Central City	Low	4
Small	Large	Large/Mid-size Central City	Medium	3
Small	Large	Large/Mid-size Central City	High	3
Small	Large	Urban Fringe of Large/Mid-size Central City	None	32
Small	Large	Large/Small Town	None	32
Small	Large	Rural	None	33
Missouri				
Small	Small	Large/Small Town	None	1
Small	Small	Rural	None	6
Small	Large	Large/Mid-size Central city	Low	4
Small	Large	Large/Mid-size Central city	Medium	4
Small	Large	Large/Mid-size Central city	High	5
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	12
Small	Large	Urban Fringe of Large/’Mid-size Central City	Medium	11
Small	Large	Urban Fringe of Large/Mid-size Central City	High	12
Small	Large	Large/Small Town	None	30
Small	Large	Rural	None	35

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Mississippi				
Small	Large	Mid-size Central City	Low	3
Small	Large	Mid-size Central City	Medium	3
Small	Large	Mid-size Central City	High	4
Small	Large	Urban Fringe of Large/'Mid-size Central City	Low	4
Small	Large	Urban Fringe of Large/'Mid-size Central City	Medium	4
Small	Large	Urban Fringe of Large/'Mid-size Central City	High	4
Small	Large	Large/Small Town	Low	15
Small	Large	Large/Small Town	Medium	15
Small	Large	Large/Small Tow	High	14
Small	Large	Rural	Low	16
Small	Large	Rural	Medium	13
Small	Large	Rural	High	14
Montana				
Small	Small	Mid-size Central City/Urban Fringe	None	1
Small	Small	Rural	None	56
Small	Large	Mid-size Central City/Urban Fringe	None	9
Small	Large	Large Town	None	6
Small	Large	Small Town	None	25
Small	Large	Rural	None	44

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
North Carolina				
Small	Large	Mid-size Central City	Low	9
Small	Large	Mid-size Central City	Medium	9
Small	Large	Mid-size Central City	High	10
Small	Large	Urban Fringe of mid-size Central City	Low	5
Small	Large	Urban Fringe of mid-size Central City	Medium	5
Small	Large	Urban Fringe of mid-size Central City	High	5
Small	Large	Large/Small Town	Low	12
Small	Large	Large/Small Town	Medium	12
Small	Large	Large/Small Town	High	12
Small	Large	Rural	Low	10
Small	Large	Rural	Medium	10
Small	Large	Rural	High	9
North Dakota				
Small	Small	Rural	None	67
Small	Large	Mid-size Central City/Urban Fringe	None	10
Small	Large	Large/Small Town	None	17
Small	Large	Rural	None	64

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Nebraska				
Small	Small	Mid-size Central City/Urban Fringe	Low	1
Small	Small	Large/Small Town	None	13
Small	Small	Rural	None	46
Small	Large	Mid-size Central City/Urban Fringe	Low	11
Small	Large	Mid-size Central City/Urban Fringe	Medium	10
Small	Large	Mid-size Central City/Urban Fringe	High	9
Small	Large	Large/Small Town	None	26
Small	Large	Rural	None	46
New Hampshire				
Small	Small	Large/Small Town	None	1
Small	Small	Rural	None	2
Small	Large	Mid-size Central City/Urban Fringe	None	14
Small	Large	Large/Small Town	None	49
Small	Large	Rural	none	23
New Jersey				
Small	Small	Large/Small Town/Rural	None	1
Small	Large	Large/Mid-size Central City	Low Black/Low Hispanic	4
Small	Large	Large/Mid-size Central City	Low Black/High Hispanic	5
Small	Large	Large/Mid-size Central City	High Black/Low Hispanic	4
Small	Large	Large/Mid-size Central City	High Black/High Hispanic	5
Small	Large	Urban Fringe of Large Central City	Low	14
Small	Large	Urban Fringe of Large Central City	Medium	14
Small	Large	Urban Fringe of Large Central City	High	14
Small	Large	Urban Fringe of Mid-size Central City	Low	7
Small	Large	Urban Fringe of Mid-size Central City	Medium	7
Small	Large	Urban Fringe of Mid-size Central City	High	7
Small	Large	Large/Small Town/Rural	None	27

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
New Mexico				
Large	Large	Mid-size Central City/Urban Fringe	Low	9
Large	Large	Mid-size Central City/Urban Fringe	Medium	6
Large	Large	Mid-size Central City/Urban Fringe	High	6
Large	Large	Rural	Medium	1
Small	Small	Rural	Low	1
Small	Small	Rural	Medium	2
Small	Small	Rural	High	1
Small	Large	Mid-size Central City/Urban Fringe	Low	1
Small	Large	Mid-size Central City/Urban Fringe	Medium	4
Small	Large	Mid-size Central City/Urban Fringe	High	5
Small	Large	Large Town	Low	5
Small	Large	Large Town	Medium	4
Small	Large	Large Town	High	4
Small	Large	Small Town	Low	9
Small	Large	Small Town	Medium	9
Small	Large	Small Town	High	8
Small	Large	Rural	Low	6
Small	Large	Rural	Medium	5
Small	Large	Rural	High	6

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Nevada				
Large	Large	Mid-size Central City	Low	3
Large	Large	Mid-size Central City	Medium	7
Large	Large	Mid-size Central City	High	7
Large	Large	Urban Fringe of Mid-size Central City	Low	1
Large	Large	Urban Fringe of Mid-size Central City	Medium	2
Large	Large	Urban Fringe of Mid-size Central City	High	2
Large	Large	Rural	Low	2
Large	Large	Rural	Medium	2
Large	Large	Rural	High	3
Small	Small	Large/Small Town	High	1
Small	Small	Rural	Low	2
Small	Small	Rural	Medium	2
Small	Small	Rural	High	1
Small	Large	Mid-size Central City	Low	5
Small	Large	Mid-size Central City	Medium	1
Small	Large	Mid-size Central City	High	1
Small	Large	Large/Small Town	Low	3
Small	Large	Large/Small Town	Medium	3
Small	Large	Large/Small Town	High	5
Small	Large	Rural	Low	5
Small	Large	Rural	Medium	3
Small	Large	Rural	High	1

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
New York				
Large	Large	Large/Mid-size Central City	Low Black/Low Hispanic	6
Large	Large	Large/Mid-size Central City	Low Black/High Hispanic	10
Large	Large	Large/Mid-size Central City	High Black/Low Hispanic	10
Large	Large	Large/Mid-size Central City	High Black/High Hispanic	7
Large	Large	Urban Fringe of Large Central City	High	1
Small	Large	Large/Mid-size Central City	Low Black/Low Hispanic	4
Small	Large	Large/Mid-size Central City	Low Black/High Hispanic	1
Small	Large	Large/Mid-size Central City	High Black/High Hispanic	4
Small	Large	Urban Fringe of Large Central City	Low	4
Small	Large	Urban Fringe of Large Central City	Medium	4
Small	Large	Urban Fringe of Large Central City	High	3
Small	Large	Urban Fringe of Mid-size Central City	Low	7
Small	Large	Urban Fringe of Mid-size Central City	Medium	6
Small	Large	Urban Fringe of Mid-size Central City	High	7
Small	Large	Large/Small Town/Rural	None	32
Oregon				
Small	Small	Urban Fringe of Large/Mid-size Central City	None	1
Small	Small	Urban Fringe of Mid-size Central City	Low	1
Small	Small	Urban Fringe of Mid-size Central City	High	1
Small	Small	Large/Small Town/Rural	None	7
Small	Large	Large Central City	Low	4
Small	Large	Large Central City	Medium	4
Small	Large	Large Central City	High	4
Small	Large	Mid-size Central City	None	12
Small	Large	Urban Fringe of Large Central City	None	33
Small	Large	Urban Fringe of Mid-size Central City	Low	4
Small	Large	Urban Fringe of Mid-size Central City	Medium	4
Small	Large	Urban Fringe of Mid-size Central City	High	5
Small	Large	Large/Small Town/Rural	None	36

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Rhode Island				
Small	Small	Large Central City	Low Hispanic/Low Black	1
Small	Small	Large/Small Town/Rural	None	1
Small	Large	Large Central City	Low Hispanic/Low Black	2
Small	Large	Large Central City	Low Hispanic/High Black	1
Small	Large	Large Central City	High Hispanic/Low Black	1
Small	Large	Large Central City	High Hispanic/High Black	2
Small	Large	Mid-size Central City	Low	2
Small	Large	Mid-size Central City	Medium	1
Small	Large	Mid-size Central City	High	2
Small	Large	Urban Fringe of Large/Mid-size Central City	None	25
Small	Large	Large/Small Town/Rural	None	14
South Carolina				
Small	Large	Mid-size Central City	Low	6
Small	Large	Mid-size Central City	Medium	5
Small	Large	Mid-size Central City	High	6
Small	Large	Urban Fringe of Mid-size Central City	Low	10
Small	Large	Urban Fringe of Mid-size Central City	Medium	9
Small	Large	Urban Fringe of Mid-size Central City	High	10
Small	Large	Small Town	Low	13
Small	Large	Small Town	Medium	14
Small	Large	Small Town	High	13
Small	Large	Rural	Low	7
Small	Large	Rural	Medium	7
Small	Large	Rural	High	7

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Tennessee				
Small	Small	Large/Small Town	Low	1
Small	Small	Rural	None	1
Small	Large	Large/Mid-size Central City	Low	10
Small	Large	Large/Mid-size Central City	Medium	11
Small	Large	Large/Mid-size Central City	High	11
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	7
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	7
Small	Large	Urban Fringe of Large/Mid-size Central City	High	7
Small	Large	Large/Small Town	Low	13
Small	Large	Large/Small Town	Medium	11
Small	Large	Large/Small Town	High	10
Small	Large	Rural	None	24

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Texas				
Small	Small	Large/Small Town	High	1
Small	Small	Rural	Low	1
Small	Small	Rural	Medium	1
Small	Large	Large Central City	Low Hispanic/Low Black	7
Small	Large	Large Central City	Low Hispanic/High Black	6
Small	Large	Large Central City	High Hispanic/Low Black	6
Small	Large	Large Central City	High Hispanic/High Black	6
Small	Large	Mid-size Central City	Low	8
Small	Large	Mid-size Central City	Medium	9
Small	Large	Mid-size Central City	High	8
Small	Large	Urban Fringe of Large/Mid-size Central City	Low	5
Small	Large	Urban Fringe of Large/Mid-size Central City	Medium	7
Small	Large	Urban Fringe of Large/Mid-size Central City	High	6
Small	Large	Large/Small Town	Low	8
Small	Large	Large/Small Town	Medium	8
Small	Large	Large/Small Town	High	8
Small	Large	Rural	Low	5
Small	Large	Rural	Medium	5
Small	Large	Rural	High	6
Utah				
Small	Small	Rural	None	1
Small	Large	Mid-size Central City	Low	9
Small	Large	Mid-size Central City	Medium	7
Small	Large	Mid-size Central City	High	8
Small	Large	Urban Fringe of Mid-size Central City	None	38
Small	Large	Large/Small Town	None	18
Small	Large	Rural	None	16

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Virginia				
Small	Large	Mid-size Central City	Low	10
Small	Large	Mid-size Central City	Medium	11
Small	Large	Mid-size Central City	High	10
Small	Large	Urban Fringe of Large Central City	None	19
Small	Large	Urban Fringe of Mid-size Central City	Low	4
Small	Large	Urban Fringe of Mid-size Central City	Medium	5
Small	Large	Urban Fringe of Mid-size Central City	High	4
Small	Large	Large/Small Town	Low	5
Small	Large	Large/Small Town	Medium	5
Small	Large	Large/Small Town	High	6
Small	Large	Rural	Low	10
Small	Large	Rural	Medium	8
Small	Large	Rural	High	9
Vermont				
Small	Small	Rural	None	19
Small	Large	-	-	2
Small	Large	Mid-size Central City/Urban Fringe	None	3
Small	Large	Small Town	None	51
Small	Large	Rural	None	41

Table F-1 (continued)

Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Washington				
Small	Small	Large/Small Town	Low	1
Small	Small	Rural	None	2
Small	Large	Large/Mid-size Central City	None	29
Small	Large	Urban Fringe of Large Central City	None	22
Small	Large	Urban Fringe of Mid-size Central City	None	16
Small	Large	Large/Small Town	Low	8
Small	Large	Large/Small Town	Medium	7
Small	Large	Large/Small Town	High	8
Small	Large	Rural	None	18
Wisconsin				
Small	Small	Rural	None	2
Small	Large	Large/Mid-size Central City	Low	10
Small	Large	Large/Mid-size Central City	Medium	11
Small	Large	Large/Mid-size Central City	High	10
Small	Large	Urban Fringe of Large/Mid-size Central City	None	18
Small	Large	Large/Small Town	None	34
Small	Large	Rural	None	30
West Virginia				
Small	Small	Rural	None	2
Small	Large	Mid-size Central City	None	15
Small	Large	Urban Fringe of Mid-size Central City	None	15
Small	Large	Large/Small Town	None	33
Small	Large	Rural	None	43

Table F-1 (continued)
Distribution of Selected Public Schools by Sampling Strata, Eighth Grade

Small or Large District	Small or Large School	Urbanization	Percent of Minority	Originally Selected Schools
Wyoming				
Small	Small	Urban Fringe of Mid-size Central City	Low	1
Small	Small	Small Town	None	8
Small	Small	Rural	None	17
Small	Large	Mid-size Central City	None	4
Small	Large	Urban Fringe of Mid-size Central City	Low	1
Small	Large	Urban Fringe of Mid-size Central City	Medium	1
Small	Large	Urban Fringe of Mid-size Central City	High	1
Small	Large	Small Town	None	28
Small	Large	Rural	None	26

Table F-2
Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Alaska			
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	19
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	4
Alabama			
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	2
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	9
Large	In Metro Area	Catholic	3
Arkansas			
Small	In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Other Nonpublic	4
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	3
Large	In Metro Area	Catholic	2
Arizona			
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	5
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	5
Large	In Metro Area	Catholic	3
California			
Small	In Metro Area	Other Nonpublic	9
Large	Not In Metro Area	Catholic	1
Large	In Metro Area	Other Nonpublic	9
Large	In Metro Area	Catholic	10
Colorado			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	7
Small	Not In Metro Area	Other Nonpublic	4
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	6
Large	In Metro Area	Catholic	5

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Connecticut			
Small	In Metro Area	Catholic	3
Small	In Metro Area	Other Nonpublic	7
Small	Not In Metro Area	Other Nonpublic	1
Large	Not In Metro Area	Catholic	2
Large	In Metro Area	Other Nonpublic	9
Large	In Metro Area	Catholic	19
District of Columbia			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	7
Large	In Metro Area	Other Nonpublic	14
Large	In Metro Area	Catholic	24
Delaware			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	21
Small	Not In Metro Area	Other Nonpublic	8
Large	In Metro Area	Other Nonpublic	13
Large	In Metro Area	Catholic	22
Florida			
Small	In Metro Area	Other Nonpublic	10
Small	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	12
Large	In Metro Area	Catholic	7
Georgia			
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	4
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	7
Large	In Metro Area	Catholic	2
Hawaii			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	5
Small	Not In Metro Area	Other Nonpublic	5
Large	Not In Metro Area	Catholic	3
Large	Not In Metro Area	Other Nonpublic	3
Large	In Metro Area	Other Nonpublic	14
Large	In Metro Area	Catholic	10

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Iowa			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	2
Small	Not In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	6
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	2
Large	In Metro Area	Catholic	8
Indiana			
Small	In Metro Area	Catholic	2
Small	In Metro Area	Other Nonpublic	11
Small	Not In Metro Area	Other Nonpublic	7
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	5
Large	In Metro Area	Catholic	10
Kentucky			
Small	In Metro Area	Catholic	2
Small	In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	4
Large	In Metro Area	Catholic	9
Louisiana			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Other Nonpublic	1
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	8
Large	In Metro Area	Catholic	20
Massachusetts			
Small	In Metro Area	Catholic	2
Small	In Metro Area	Other Nonpublic	7
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	8
Large	In Metro Area	Catholic	20

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Maryland			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	8
Small	Not In Metro Area	Other Nonpublic	1
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	11
Large	In Metro Area	Catholic	15
Maine			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Other Nonpublic	13
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	4
Large	In Metro Area	Catholic	3
Michigan			
Small	In Metro Area	Catholic	2
Small	In Metro Area	Other Nonpublic	10
Small	Not In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	1
Large	In Metro Area	Other Nonpublic	8
Large	In Metro Area	Catholic	10
Minnesota			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	8
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	4
Large	In Metro Area	Catholic	10

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Missouri			
Small	In Metro Area	Catholic	3
Small	In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	6
Large	In Metro Area	Catholic	17
Mississippi			
Small	In Metro Area	Other Nonpublic	2
Small	Not In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	10
Large	In Metro Area	Other Nonpublic	3
Large	In Metro Area	Catholic	1
Montana			
Small	In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Other Nonpublic	16
Small	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Catholic	3
Large	Not In Metro Area	Other Nonpublic	4
Large	In Metro Area	Other Nonpublic	1
Large	In Metro Area	Catholic	1
North Carolina			
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	5
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	6
Large	In Metro Area	Catholic	2
North Dakota			
Small	In Metro Area	Other Nonpublic	5
Small	Not In Metro Area	Other Nonpublic	12
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	2
Large	In Metro Area	Catholic	5

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Nebraska			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Other Nonpublic	10
Small	Not In Metro Area	Catholic	3
Large	Not In Metro Area	Catholic	6
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	4
Large	In Metro Area	Catholic	10
New Hampshire			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	9
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	2
Large	In Metro Area	Other Nonpublic	2
Large	In Metro Area	Catholic	7
New Jersey			
Small	In Metro Area	Catholic	4
Small	In Metro Area	Other Nonpublic	6
Large	In Metro Area	Other Nonpublic	9
Large	In Metro Area	Catholic	28
New Mexico			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	9
Small	Not In Metro Area	Other Nonpublic	7
Small	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	3
Large	In Metro Area	Other Nonpublic	3
Large	In Metro Area	Catholic	5
Nevada			
Small	In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Other Nonpublic	2
Large	Not In Metro Area	Catholic	1
Large	In Metro Area	Other Nonpublic	3
Large	In Metro Area	Catholic	4

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
New York			
Small	In Metro Area	Catholic	2
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	2
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	12
Large	In Metro Area	Catholic	21
Oregon			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	7
Small	Not In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Catholic	1
Large	In Metro Area	Other Nonpublic	5
Large	In Metro Area	Catholic	5
Rhode Island			
Small	In Metro Area	Catholic	5
Small	In Metro Area	Other Nonpublic	8
Small	Not In Metro Area	Other Nonpublic	1
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	1
Large	In Metro Area	Other Nonpublic	4
Large	In Metro Area	Catholic	26
South Carolina			
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	4
Large	Not In Metro Area	Other Nonpublic	3
Large	In Metro Area	Other Nonpublic	6
Large	In Metro Area	Catholic	2
Tennessee			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	6
Small	Not In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	10
Large	In Metro Area	Catholic	3

Table F-2 (continued)
Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Texas			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	6
Large	In Metro Area	Catholic	3
Utah			
Small	In Metro Area	Other Nonpublic	4
Small	Not In Metro Area	Other Nonpublic	1
Small	Not In Metro Area	Catholic	1
Large	In Metro Area	Other Nonpublic	5
Large	In Metro Area	Catholic	2
Virginia			
Small	In Metro Area	Other Nonpublic	7
Small	Not In Metro Area	Other Nonpublic	4
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	7
Large	In Metro Area	Catholic	4
Vermont			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Other Nonpublic	19
Small	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Catholic	2
Large	Not In Metro Area	Other Nonpublic	5
Large	In Metro Area	Other Nonpublic	1
Large	In Metro Area	Catholic	3
Washington			
Small	In Metro Area	Catholic	1
Small	In Metro Area	Other Nonpublic	7
Small	Not In Metro Area	Other Nonpublic	3
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	7
Large	In Metro Area	Catholic	6

Table F-2 (continued)

Distribution of Selected Nonpublic Schools by Sampling Strata, Eighth Grade

Small or Large School	Metro Status	School Type	Original Selected School
Wisconsin			
Small	In Metro Area	Catholic	6
Small	In Metro Area	Other Nonpublic	14
Small	Not In Metro Area	Other Nonpublic	11
Small	Not In Metro Area	Catholic	5
Large	Not In Metro Area	Catholic	4
Large	Not In Metro Area	Other Nonpublic	3
Large	In Metro Area	Other Nonpublic	8
Large	In Metro Area	Catholic	17
West Virginia			
Small	In Metro Area	Catholic	2
Small	In Metro Area	Other Nonpublic	5
Small	Not In Metro Area	Other Nonpublic	11
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Other Nonpublic	1
Large	In Metro Area	Catholic	3
Wyoming			
Small	In Metro Area	Other Nonpublic	3
Small	Not In Metro Area	Other Nonpublic	11
Small	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Catholic	1
Large	Not In Metro Area	Other Nonpublic	1
Large	In Metro Area	Catholic	2

Table F-3

Number of Substitute Schools Provided in Each Jurisdiction By School Type, Fourth Grade

Jurisdiction	Public			Nonpublic		
	Regular substitutes	**Double session substitutes**	**Total substitutes**	**Regular substitutes**	**Double session substitutes**	**Total substitutes**
DoDEA/DDESS	0	0	0	N/A	N/A	N/A
DoDEA/DoDDS	0	0	0	N/A	N/A	N/A
Total	**0**	**0**	**0**	**N/A**	**N/A**	**N/A**

Table F-4

Number of Substitute Schools Provided in Each Jurisdiction By School Type, Eighth Grade

Jurisdiction	Public			Nonpublic		
	Regular substitutes	Double session substitutes	Total substitutes	Regular substitutes	Double session substitutes	Total substitutes
Alabama	89	2	91	19	0	19
Alaska	91	0	91	0	0	0
Arizona	59	0	59	0	0	0
Arkansas	74	0	74	12	0	12
California	97	0	97	28	0	28
Colorado	65	0	65	0	0	0
Connecticut	46	0	46	37	0	37
Delaware	0	0	0	30	0	30
District of Columbia	0	0	0	0	0	0
DoDEA/DDESS	0	0	0	N/A	N/A	N/A
DoDEA/DoDDS	0	0	0	N/A	N/A	N/A
Florida	86	0	86	0	0	0
Georgia	96	0	96	20	0	20
Guam	0	0	0	0	0	0
Hawaii	0	0	0	0	0	0
Indiana	93	0	93	0	0	0
Iowa	92	0	92	27	0	27
Kentucky	85	0	85	24	0	24
Louisiana	90	0	90	31	0	31
Maine	72	0	72	0	0	0
Maryland	66	0	66	37	0	37
Massachusetts	75	0	75	35	0	35
Michigan	102	0	102	35	0	35
Minnesota	75	0	75	30	0	30
Mississippi	70	0	70	0	0	0
Missouri	100	0	100	38	0	38
Montana	103	0	103	28	0	28

Table F-4 (continued)

Number of Substitute Schools Provided in Each Jurisdiction By School Type, Eighth Grade

Jurisdiction	Public			Nonpublic		
	Regular substitutes	Double session substitutes	Total substitutes	Regular substitutes	Double session substitutes	Total substitutes
Nebraska	118	0	118	37	0	37
Nevada	12	0	12	8	0	8
New Hampshire	36	0	36	23	0	23
New Jersey	93	0	93	39	0	39
New Mexico	19	0	19	24	0	24
New York	96	0	96	38	0	38
North Carolina	98	0	98	0	0	0
North Dakota	79	4	83	17	0	17
Oregon	82	0	82	22	0	22
Rhode Island	0	4	4	21	0	21
South Carolina	66	0	66	19	0	19
Tennessee	94	0	94	0	0	0
Texas	103	0	103	12	0	12
Utah	14	0	14	11	0	11
Vermont	12	0	12	16	0	16
Virginia	82	0	82	0	0	0
Washington	105	0	105	21	0	21
West Virginia	54	0	54	0	0	0
Wisconsin	101	0	101	64	0	64
Wyoming	10	0	10	9	0	9
Total	**3,000**	**10**	**3,010**	**812**	**0**	**812**

Table F-5

Distribution of Grade 4 Public-School Sample by Jurisdiction

Jurisdiction	Weighted percent of school participation		Number of schools in the original sample			Number of substitute schools for nonparticipating originals		Total number of schools that participated
	Before substitution	After substitution	Total	Not eligible	Participated	Provided	Participated	
DoDEA/DDESS	100%	100%	39	0	39	0	0	39
DoDEA/DoDDS	100%	100%	92	1	91	0	0	91

Table F-6

Distribution of Grade 8 Public-School Sample by Jurisdiction

Jurisdiction	Weighted percent of school participation		Number of schools in the original sample			Number of substitute schools for nonparticipating originals		Total number of schools that participated
	Before substitution	After substitution	Total	Not eligible	Participated	Provided	Participated	
Alabama	84%	90%	109	2	90	15	6	96
Alaska	93%	93%	81	5	55	17	0	55
Arizona	87%	87%	108	1	94	9	0	94
Arkansas	70%	71%	111	1	75	25	1	76
California	83%	94%	108	0	89	19	12	101
Colorado	100%	100%	110	2	108	0	0	108
Connecticut	100%	100%	103	1	102	0	0	102
Delaware	100%	100%	31	1	30	0	0	30
District of Columbia	100%	100%	36	3	33	0	0	33
DoDEA/DDESS	100%	100%	11	0	11	0	0	11
DoDEA/DoDDS	100%	100%	59	1	58	0	0	58
Florida	100%	100%	110	5	105	0	0	105
Georgia	99%	99%	107	6	100	1	0	100
Guam	100%	100%	6	0	6	0	0	6
Hawaii	100%	100%	52	1	51	0	0	51
Indiana	87%	90%	107	1	93	12	3	96
Iowa	73%	83%	115	4	80	27	11	91
Kentucky	87%	92%	109	1	95	9	5	100
Louisiana	100%	100%	114	3	111	0	0	111
Maine	91%	91%	110	4	95	6	0	95
Maryland	86%	86%	106	2	89	10	0	89
Massachusetts	92%	92%	107	2	98	7	0	98
Michigan	70%	87%	107	1	74	31	18	92
Minnesota	86%	88%	107	0	93	8	2	95
Mississippi	89%	95%	109	3	96	9	7	103
Missouri	93%	96%	116	6	102	7	3	105

Table F-6 (continued)

Distribution of Grade 8 Public-School Sample by Jurisdiction

Jurisdiction	Weighted percent of school participation		Number of schools in the original sample			Number of substitute schools for nonparticipating originals		Total number of schools that participated
	Before substitution	After substitution	Total	Not eligible	Participated	Provided	Participated	
Montana	70%	76%	113	5	68	28	11	79
Nebraska	99%	100%	132	12	119	1	1	120
Nevada	37%	38%	59	2	27	3	1	28
New Hampshire	66%	68%	88	0	61	10	3	64
New Jersey	63%	64%	108	2	66	35	1	67
New Mexico	100%	100%	90	0	90	0	0	90
New York	70%	78%	106	0	74	30	8	82
North Carolina	100%	100%	108	1	107	0	0	107
North Dakota	80%	93%	125	7	94	19	14	108
Oregon	86%	92%	111	3	94	12	6	100
Rhode Island	90%	90%	51	0	43	4	0	43
South Carolina	86%	87%	107	2	90	10	1	91
Tennessee	92%	92%	112	4	99	6	0	99
Texas	91%	96%	109	3	97	9	5	102
Utah	100%	100%	96	2	94	0	0	94
Vermont	74%	75%	104	5	77	1	1	78
Virginia	100%	100%	106	0	106	0	0	106
Washington	94%	95%	109	0	104	5	1	105
West Virginia	100%	100%	107	2	105	0	0	105
Wisconsin	78%	78%	114	0	90	23	0	90
Wyoming	100%	100%	74	7	67	0	0	67

Table F-7

Distribution of Grade 8 Nonpublic-School Sample by Jurisdiction

Jurisdiction	Weighted percent of school participation		Number of schools in the original sample			Number of substitute schools for nonparticipating originals		Total number of schools that participated
	Before substitution	After substitution	Total	Not eligible	Participated	Provided	Participated	
Alabama	60%	60%	19	2	10	7	0	10
Arkansas	74%	74%	11	2	6	2	0	6
California	80%	80%	25	6	14	5	0	14
Connecticut	63%	65%	36	6	19	10	1	20
Delaware	42%	44%	50	14	12	6	1	13
District of Columbia	52%	52%	42	8	19	0	0	19
Georgia	88%	88%	16	5	9	2	0	9
Guam	79%	79%	10	0	8	0	0	8
Iowa	94%	94%	22	6	14	2	0	14
Kentucky	82%	82%	21	4	13	4	0	13
Louisiana	75%	75%	35	5	21	7	0	21
Maryland	61%	64%	34	3	18	12	1	19
Massachusetts	75%	77%	33	5	20	8	1	21
Michigan	80%	87%	28	3	19	6	2	21
Minnesota	84%	84%	25	3	19	3	0	19
Missouri	94%	100%	33	9	23	1	1	24

Table F-7 (continued)

Distribution of Grade 8 Nonpublic-School Sample by Jurisdiction

Jurisdiction	Weighted percent of school participation		Number of schools in the original sample			Number of substitute schools for nonparticipating originals		Total number of schools that participated
	Before substitution	After substitution	Total	Not eligible	Participated	Provided	Participated	
Montana	93%	97%	20	5	12	3	1	13
Nebraska	78%	84%	31	6	17	7	3	20
Nevada	90%	90%	10	1	8	1	0	8
New Hampshire	83%	83%	20	4	12	3	0	12
New Jersey	62%	64%	42	10	20	11	0	20
New Mexico	95%	95%	21	6	13	2	0	13
New York	84%	87%	39	6	27	5	1	28
North Dakota	70%	78%	20	7	9	2	1	10
Oregon	26%	26%	17	4	4	8	0	4
Rhode Island	68%	68%	38	5	22	5	0	22
South Carolina	69%	69%	16	4	8	3	0	8
Texas	79%	79%	11	1	7	3	0	7
Utah	64%	64%	10	1	4	3	0	4
Vermont	72%	80%	23	10	9	2	1	10
Washington	86%	86%	20	5	11	3	0	11
Wisconsin	65%	69%	50	8	25	15	2	27
Wyoming	92%	92%	11	4	6	1	0	6

Table F-8
Distribution of the Fourth-Grade Public-School Student Sample and Response Rates by Jurisdiction

Jurisdiction	Weighted Percentage of Student Participation After Make-Ups	Number of Students Original Sample	Number of Students Excluded	Number of Students to be Assessed	Total Number of Students Assessed
DoDEA/DDESS	96%	1,404	41	1,310	1,251
DoDEA/DoDDS	94%	2,948	60	2,718	2,567

Table F-9

Distribution of the Eighth-Grade Public-School Student Sample and Response Rates by Jurisdiction

Jurisdiction	Weighted Percentage of Student Participation After Make-Ups	Number of Students Original Sample	Number of Students Excluded	Number of Students to be Assessed	Total Number of Students Assessed
Alabama	93%	2,550	99	2,356	2,186
Alaska	82%	1,931	45	1,821	1,517
Arizona	90%	2,560	88	2,402	2,151
Arkansas	92%	2,120	75	2,010	1,858
California	92%	2,664	131	2,493	2,292
Colorado	91%	2,955	105	2,745	2,514
Connecticut	93%	2,887	141	2,696	2,489
Delaware	89%	2,189	32	2,127	1,903
District of Columbia	85%	2,186	98	2,005	1,700
DoDEA/DDESS	95%	686	23	635	602
DoDEA/DoDDS	93%	2,562	41	2,376	2,223
Florida	90%	2,812	153	2,613	2,353
Georgia	92%	2,833	81	2,692	2,470
Guam	90%	1,077	47	1,033	930
Hawaii	90%	2,565	61	2,394	2,153
Indiana	92%	2,665	87	2,516	2,313
Iowa	94%	2,417	72	2,320	2,172
Kentucky	94%	2,701	60	2,598	2,459
Louisiana	90%	3,046	94	2,894	2,615
Maine	92%	2,559	105	2,443	2,254
Maryland	89%	2,482	76	2,350	2,092
Massachusetts	91%	2,640	118	2,489	2,287
Michigan	90%	2,547	74	2,429	2,186
Minnesota	92%	2,699	57	2,597	2,383
Mississippi	92%	2,860	88	2,693	2,469
Missouri	92%	2,746	89	2,607	2,389

Table F-9 (continued)

Distribution of the Eighth-Grade Public-School Student Sample and Response Rates by Jurisdiction

Jurisdiction	Weighted Percentage of Student Participation After Make-Ups	Number of Students Original Sample	Number of Students Excluded	Number of Students to be Assessed	Total Number of Students Assessed
Montana	92%	2,305	47	2,206	2,029
Nebraska	92%	3,070	65	2,980	2,724
Nevada	92%	1,112	34	1,048	964
New Hampshire	90%	1,975	58	1,898	1,710
New Jersey	93%	1,827	109	1,685	1,573
New Mexico	90%	2,870	143	2,628	2,377
New York	90%	2,204	113	2,068	1,876
North Carolina	91%	2,981	77	2,861	2,616
North Dakota	94%	2,692	18	2,641	2,489
Oregon	89%	2,718	87	2,555	2,275
Rhode Island	89%	2,482	93	2,344	2,087
South Carolina	90%	2,523	79	2,386	2,162
Tennessee	91%	2,631	52	2,500	2,287
Texas	92%	2,701	126	2,497	2,300
Utah	90%	3,122	74	2,993	2,715
Vermont	93%	2,149	66	2,050	1,914
Virginia	90%	2,975	112	2,805	2,552
Washington	90%	2,871	58	2,767	2,501
West Virginia	93%	2,984	110	2,801	2,602
Wisconsin	90%	2,506	106	2,370	2,148
Wyoming	93%	2,932	64	2,796	2,619

Table F-10

Distribution of the Eighth-Grade Nonpublic-School Student Sample and Response Rates by Jurisdiction

Jurisdiction	Weighted Percentage of Student Participation After Make-Ups	Number of Students Original Sample	Number of Students Excluded	Number of Students to be Assessed	Total Number of Students Assessed
Alabama	95%	151	0	150	144
Arkansas	99%	90	0	90	89
California	96%	213	0	214	206
Connecticut	96%	278	2	274	263
Delaware	96%	325	0	325	313
District of Columbia	95%	276	0	274	259
Georgia	96%	244	0	242	232
Guam	94%	213	0	210	198
Iowa	96%	255	0	256	246
Kentucky	97%	269	0	268	260
Louisiana	96%	448	1	445	424
Maryland	94%	347	0	343	322
Massachusetts	94%	360	3	357	335
Michigan	97%	344	3	344	332
Minnesota	94%	266	0	264	247
Missouri	95%	383	1	385	365
Montana	93%	158	1	163	154
Nebraska	97%	346	1	346	333
Nevada	91%	149	2	146	133
New Hampshire	95%	189	0	188	179
New Jersey	97%	299	1	297	287
New Mexico	95%	243	0	241	230
New York	97%	537	3	530	514
North Dakota	93%	169	1	168	160

Table F-10 (continued)

Distribution of the Eighth-Grade Nonpublic-School Student Sample and Response Rates by Jurisdiction

Jurisdiction	Weighted Percentage of Student Participation After Make-Ups	Number of Students Original Sample	Number of Students Excluded	Number of Students to be Assessed	Total Number of Students Assessed
Oregon	86%	62	0	62	54
Rhode Island	96%	359	3	352	340
South Carolina	95%	142	0	144	138
Texas	98%	132	0	133	130
Utah	93%	99	0	100	93
Vermont	91%	131	1	127	115
Washington	95%	235	0	227	215
Wisconsin	96%	395	1	393	380
Wyoming	94%	47	0	50	47

REFERENCES CITED IN TEXT

Abedi, J. (1996). The Interrater/Test Reliability System (ITRS). *Multivariate Behavioral Research*, 31(4), 409-417.

Allen, N. L., Carlson, J. E., & Zelenak, C. A. (1998). *The NAEP 1996 technical report*. Washington, DC: National Center for Education Statistics. Report in preparation.

Allen, N. L., Jenkins, F., Kulick, E., & Zelenak, C.A. (1997). *Technical report of the NAEP 1996 state assessment program in mathematics*. Washington, DC: National Center for Education Statistics.

Allen, N. L., Kline, D. L., & Zelenak, C. A. (1996). *The NAEP 1994 technical report*. Washington, DC: National Center for Education Statistics.

Allen, N. L., & Liang, J. (1997, April). *Relationships between test specifications, item response, and item attributes in the 1996 NAEP science assessment*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

Andersen, E. B. (1980). Comparing latent distributions. *Psychometrika, 45*, 121-134.

Anderson, N. E., Jenkins, F. F., & Miller, K. E. (1996). *NAEP inclusion criteria and testing accommodations: Findings from the NAEP 1995 field test in mathematics*. Princeton, NJ: Educational Testing Service.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508 - 600). Washington, DC: American Council on Education.

Ballator, N., O'Sullivan, C. Y., & Jerry, L. (1997a). *NAEP 1996 Science Report for Department of Defense Dependents Schools, Grade 4*. Washington, DC: National Center for Education Statistics.

Ballator, N., O'Sullivan, C. Y., & Jerry, L. (1997b). *NAEP 1996 Science Report for Domestic Dependent Elementary and Secondary Schools, Grade 4*. Washington, DC: National Center for Education Statistics.

Beaton, A. E., & Johnson, E. G. (1990). The average response method of scaling. *Journal of Educational Statistics*, *15,* 9-38.

Beaton, A. E., & Johnson, E. G. (1992). Overview of the scaling methodology used in the National Assessment. *Journal of Educational Measurement*, *26(2)*, 163-175.

Beaton, A. E., & Zwick, R. (1990). *The effect of changes in the National Assessment: Disentangling the NAEP 1985-86 reading anomaly*. (No. 17-TR-21) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Benjamini, Y., & Hochberg, Y. (1994). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57*(1), 289-300.

Bourque, M. L., & Garrison, H. H. (1991). *The levels of mathematics achievement. Vol. I, national and state summaries*. Washington, DC: National Assessment Governing Board.

Carlson, J. E. (1996, April). *A comparison of information available from dichotomous and polytomous items*. Paper presented at the annual meeting of the American Educational Research Association, New York City, NY.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

Curry, L. (1987, April). *Group decision process in setting cut-off scores.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B, 39,* 1-38.

Educational Testing Service (1987). *ETS standards for quality and fairness*. Princeton, NJ: Author.

Educational Testing Service (1992). *Innovations and ingenuity: A foundation for the future. Application for cooperative agreement for NAEP*. CFDA Number: 84.999E. Princeton, NJ: Author.

Engelen, R. J. H. (1987). *Semiparametric estimation in the Rasch model*. Research Report 87-1. Twente, the Netherlands: Department of Education, University of Twente.

Fitzpatrick, A. R. (1989). Social influences in standard-setting: The effects of social interaction on group judgments. *Review of Educational Research*, *59*, 315-328.

Friedman, C. B., & Ho, K. T. (1990, April). *Interjudge consensus and intrajudge consistency: Is it possible to have both on standard setting?* Paper presented at the annual meeting of the National Council for Measurement in Education, Boston, MA.

Hambleton, R. K., & Bourque, M. L. (1991). *The levels of mathematics achievement. Vol. II, technical report.* Washington, DC: National Assessment Governing Board.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75,* 800-802.

Hoijtink, H. (1991). *Estimating the parameters of linear models with a latent dependent variable by nonparametric maximum likelihood*. Research Bulletin HB-91-1040-EX. Groningen, The Netherlands: Psychological Institute, University of Groningen.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity*. Hillsdale, NJ: Erlbaum.

Jerry, L. (1995). *The NAEP computer-generated reporting system for the 1994 Trial State Assessment*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress. Unpublished documentation.

Johnson, E. G. (1994, April). *Special studies on constructed response scoring*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Johnson, E. G., & Rust, K. F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics, 17*, 175-190.

Kaplan, B. A., & Johnson, E. G. (1992, April). *Reliability of professionally scored data: NAEP-related issues*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Keyfitz, N. (1951). Sampling with probability proportional to size; adjustment for changes in probabilities. *Journal of the American Statistical Association*, *46,* 105-109.

Laird, N. M. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association, 73*, 805-811.

Lindsey, B., Clogg, C. C., & Grego, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *Journal of the American Statistical Association, 86*, 96-107.

Little, R. J. A., & Rubin, D. B. (1983). On jointly estimating parameters and missing data. *American Statistician*, *37,* 218-220.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York, NY: John Wiley & Sons.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association, 58*, 690-700.

Mazzeo, J. (1991). Data analysis and scaling. In S. L. Koffler, *The technical report of NAEP's 1990 Trial State Assessment program* (No. ST-21-01). Washington, DC: National Center for Education Statistics.

Mazzeo, J., Allen, N. L., & Kline, D. L. (1995). *Technical report of the NAEP 1994 Trial State Assessment program in reading*. Washington, DC: National Center for Education Statistics.

Mazzeo, J., Johnson, E. G., Bowker, D., & Fong, Y. F. (1992). *The use of collateral information in proficiency estimation for the Trial State Assessment*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Miller, R. G. (1966). *Simultaneous statistical inference*. New York, NY: McGraw-Hill.

Mislevy, R. J. (1985). Estimation of latent group effects. *Journal of the American Statistical Association, 80*, 993-997.

Mislevy, R. J. (1990). Scaling procedures. In E. G. Johnson and R. Zwick, *Focusing the new design: The NAEP 1988 technical report* (No. 19-TR-20). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56*, 177-196.

Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement, 29*(2), 133-161.

Mislevy, R. J., & Bock, R. D. (1982). *BILOG: Item analysis and test scoring with binary logistic models* [Computer program]. Mooresville, IN: Scientific Software.

Mislevy, R. J., Johnson, E. G., & Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Educational Statistics, 17*(2), 131-154.

Mislevy, R. J., & Sheehan, K. M. (1987). Marginal estimation procedures. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (No. 15-TR-20). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Mislevy, R. J., & Stocking, M. L. (1987). *A consumer's guide to LOGIST and BILOG.* (ETS Research Report 87-43). Princeton, NJ: Educational Testing Service.

Mislevy, R. J., & Wu, P-K. (1988). *Inferring examinee ability when some item responses are missing* (ETS Research Report RR-88-48-ONR). Princeton, NJ: Educational Testing Service.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159-176.

Muraki, E., & Bock, R. D. (1991). *PARSCALE: Parameter scaling of rating data*. Chicago, IL: Scientific Software, Inc.

National Assessment Governing Board (1990). *Setting appropriate achievement levels for the National Assessment of Educational Progress: Policy framework and technical procedures.* Washington, DC: Author.

National Assessment Governing Board (1993). *Science framework for the 1996 National Assessment of Educational Progress*. Washington, DC: Author.

National Council of Teachers of Mathematics (1987). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Olson, J. F., & Goldstein, A. A. (1997). *The inclusion of students with disabilities and limited English proficient students in large-scale assessments*. Washington, DC: National Center for Education Statistics.

O'Reilly, P. O., Zelenak, C. A., Rogers, A. M., & Kline, D. L. (1997). *NAEP 1994 secondary-use data files user guide*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress. Report in preparation.

O'Reilly, P. O., Zelenak, C. A., Rogers, A. M., & Kline, D. L. (1997). *NAEP 1996 state assessment program in science secondary-use data files user guide*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress. Report in preparation.

O'Sullivan, C. Y., Reese, C. M., & Mazzeo, J. (1997). *NAEP 1996 science report card for the nation and the states: Findings from the National Assessment of Educational Progress*. Washington, DC: National Center for Education Statistics.

Rogers, A. M. (1991). *NAEP-MGROUP: Enhanced version of Sheehan's software for the estimation of group effects in multivariate models* [Computer program]. Princeton, NJ: Educational Testing Service.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley & Sons.

Rubin, D. B. (1991). EM and beyond. *Psychometrika, 56*, 241-254.

Rust, K. R., & Johnson, E. G. (1992). Sampling and weighting in the national assessment. *Journal of Educational Statistics, 17*(2), 111-129.

Sheehan, K. M. (1985). *MGROUP: Estimation of group effects in multivariate models* [Computer program]. Princeton, NJ: Educational Testing Service.

Smith, R. L., & Smith, J. K. (1988). Differential use of item information by judges using Angoff and Nedelsky procedures. *Journal of Educational Measurement*, *25*, 259-274.

Somes, G. W. (1986). The generalized Mantel-Haenszel statistic. *The American Statistician, 40*, 106-108.

Stone, C. A., Mislevy, R. J., & Mazzeo, J. (1994, April). *Misclassification error and goodness-of-fit in IRT models*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Tatsuoka, K. K. (1997, April). *Application of the rule space model to the 1996 NAEP science assessment*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

Tanner, M., & Wong, W. (1987). The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association, 82*, 528-550.

Thomas, N. (1993). Asymptotic corrections for multivariate posterior moments with factored likelihood functions. *Journal of Computational and Graphical Statistics*, *2*, 309-322.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley Publishing Co.

Wainer, H. (1974). The suspended rootogram and other visual displays: An empirical validation. *The American Statistician, 28(4)*, 143-145.

Wainer, H., & Kiely, G. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement, 24*, 185-202.

Williams, P. L., Reese, C. M., Campbell, J. R., Mazzeo, J., & Phillips, G. W. (1995). *NAEP 1994 reading: A first look Findings from the National Assessment of Educational Progress*. Washington, DC: National Center for Education Statistics.

Westat, Inc. (1996). *Report on the data collection activities for the 1996 National Assessment of Educational Progress*. Rockville, MD: Author.

Westat, Inc. (1996). *1996 NAEP state assessment report on data collection activities for all states*. Rockville, MD: Author.

Wingersky, M., Kaplan, B. A., & Beaton, A. E. (1987). Joint estimation procedures. In A. E. Beaton, *Implementing the new design: The NAEP 1983-84 technical report* (pp. 285-292). (No. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Worthington, L., & Donoghue, J. R. (1997, April). *Detecting and describing local item dependence using IRT methods*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Yamamoto, K., & Jenkins, F. (1992). Data analysis for the mathematics assessment. In E. G. Johnson & N. L. Allen, *The NAEP 1990 technical report* (No. 21-TR-20). Washington, DC: National Center for Education Statistics.

Yamamoto, K., & Mazzeo, J. (1992). Item response theory scale linking in NAEP. *Journal of Educational Statistics*, *17(2)*, 155-173.

Yepes-Baraya, M. (1997, April). *The coding of item attributes for the 1996 NAEP science assessment: Lessons learned from discrepancies between raters*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

Yepes-Baraya, M., & Allen, N. L. (1996, April). *A cognitive analysis based on the NAEP science assessment*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York City, NY.

Zieky, M. (1993). Practical questions in the use of DIF statistics. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Zwick, R. (1988). *Professionally scored items*. Technical memorandum.

Zwinderman, A. H. (1991). Logistic regression Rasch models. *Psychometrika, 56*, 589-600.

NCES 98–480