Education Longitudinal Study of 2002: Base-Year to First Follow-up Data File Documentation

U.S. Department of Education
Institute of Education Sciences
NCES 2006–344

National Center for Education Statistics

Education Longitudinal Study of 2002: Base-Year to First Follow-up Data File Documentation

U.S. Department of Education
Institute of Education Sciences
NCES 2006–344

October 2005

Steven J. Ingels
Daniel J. Pratt
James E. Rogers
Peter H. Siegel
Ellen S. Stutts
RTI International

Jeffrey A. Owings
Project Officer
National Center for Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

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

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

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

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

October 2005

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

Suggested Citation

Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E.S. (2005). *Education Longitudinal Study of 2002: Base-Year to First Follow-up Data File Documentation* (NCES 2006–344). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

For ordering information on this report, write to

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

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

Content Contact:

Jeffrey A. Owings
(202) 502-7423
Jeffrey.Owings@ed.gov

Foreword

This manual has been produced to familiarize data users with the procedures followed for data collection and processing for the base year and first follow-up of the Education Longitudinal Study of 2002 (ELS:2002). It also provides the necessary documentation for use of the public-use data files, as they appear on the ELS:2002 base-year to first follow-up electronic codebook (ECB) (NCES 2006-346).

Analysts do not need to be sophisticated statisticians or computer programmers to use the ELS:2002 ECB. Most social scientists and policy analysts should find the dataset organized and equipped in a manner that facilitates straightforward production of statistical summaries and analyses. This manual provides extensive documentation of the content of the data files and how to access and manipulate them.

Chapter 1 serves as an introduction to ELS:2002. It includes an overview and history of the National Center for Education Statistics (NCES) program of longitudinal high school cohorts, summarizes the ELS:2002 objectives, and supplies an overview of the base-year and longitudinal study design.

Chapter 2 describes the first follow-up data collection instruments, including both the development and content of the in-school student, transfer, dropout, early graduate, homeschooled, and school administrator questionnaires, as well as the student assessment in mathematics.

The sample design and weighting procedures used both in the base-year and first follow-up studies are documented in chapter 3, as are weights, imputation, and the calculation of design effects.

Data collection schedules, training, procedures, and results are presented in chapter 4. Chapter 5 describes data preparation and processing, including the receipt control system, optical scanning, machine editing, and data file preparation. Chapter 6 describes the contents of the data files, including the data structure and analysis populations.

The appendixes include, among other topics, an introduction to the public-use ECB (appendix A), bse-year and first follow-up questionnaires (appendix B), documentation for imputed variables (appendix C), information on variables not included in the public-use files but available in restricted files for licensed users (appendix D), a glossary of terms (appendix E), student questionnaire critical items (appendix F), cross-cohort comparisons (appendix H), and a synopsis of the ELS:2002 first follow-up field test (appendix J).

Jeffrey A. Owings
Associate Commissioner
Elementary/Secondary & Libraries Studies

Acknowledgments

Daniel J. Pratt of RTI served as the ELS:2002 base-year and first follow-up project director. Steven J. Ingels of RTI was principal investigator. Jeffrey A. Owings served as the NCES project officer. Key RTI task leaders were Ellen Stutts (associate project director), Debbie Herget (first follow-up in-school data collection), Doug Currivan (first follow-up out-of-school data collection), James Rogers (data processing), and Peter Siegel (sampling and statistics). Other RTI staff who played major roles in ELS:2002 were Christopher Alexander, Kimberly Ault, Stephen Black, Laura J. Burns, Debbie Capps, James Chromy, Elizabeth Copello, Marianne Daye, D. Wesley Dukes, Brian Evans, Catherine Forstner, Sherry Hubbard-Bednasz, Ruby Johnson, Tiffany Lytle, Mani Medarametla, Greg Mosorjak, Melanie Pressley, Denise Rhatigan, Helen Smith, Milorad Stojanovic, David Wilson, and Donghui Wang. Assessment development, scaling, and equating were conducted by Judith M. Pollack, Donald A. Rock, and Michelle Najarian, under a subcontract with Educational Testing Service (ETS). ETS staff contributed assessment documentation to this manual.

The authors of this report would like to thank the many individuals who assisted in the planning of ELS:2002. We are particularly indebted to the ELS:2002 Technical Review Panel, whose members reviewed plans for the study, helped refine them, and provided important suggestions to help guide development of the instrumentation. The following individuals serve as members of the ELS:2002 Technical Review Panel: Marco Clark, Richard Duran, Jeremy Finn, Thomas B. Hoffer, Thomas Kane, Sally Kilgore, Richard Lawrence, Samuel R. Lucas, Aaron Pallas, and Andy Rogers.

Special thanks are in order to Jeffrey A. Owings, Associate Commissioner for Elementary/Secondary & Libraries Studies at NCES, in whose division ELS:2002 is housed. Other NCES staff who have provided help, support, and assistance include Lisa Hudson, Steve Kaufman, Mariann Lemke, Andrew G. Malizio, Edith McArthur, and Marilyn M. Seastrom. Staff in other offices of the U.S. Department of Education who have contributed to the study include Clifford Adelman, Sharon Belli, Jeffery Rodamar, and Marsha Silverberg.

Special thanks are also due to Leslie A. Scott of the Education Statistics Services Institute and to Denise M. Davis, formerly of the National Commission on Libraries and Information Science.

Many others, far too many to name individually, contributed to ELS:2002, including survey administrators and other contractor and subcontractor staff, and of course the literally thousands of students, parents, and school personnel who generously gave of their time and approval to provide the data that the study reports. We extend our thanks to all.

Lastly, we want to acknowledge the help of several other individuals at RTI in the preparation of this document: Michael Planty, who assisted in the review process; Wallace Campbell and Sallie Fiore, who edited the document; Lawanda King and Sharon Powell, who provided assistance in document production; and Diane Caudill, who provided graphics support.

Contents

Page

Page

List of Tables

Table Page

List of Appendix Tables

List of Figures

Chapter 1
Introduction

1.1 Overview of the Data File Documentation

This report provides guidance and documentation for users of the public release for the combined base-year and first follow-up data of the Education Longitudinal Study of 2002 (ELS:2002). ELS:2002 is sponsored by the National Center for Education Statistics (NCES) of the Institute of Education Sciences, U.S. Department of Education. The base-year and first follow-up studies were conducted through a contract to RTI International (RTI),[1] a university-affiliated, nonprofit research organization based in North Carolina, in collaboration with its subcontractors, the Educational Testing Service of Princeton, New Jersey, and MPR Associates of Berkeley, California. This manual contains information about the purposes of ELS:2002, the base-year and first follow-up data collection instruments, the sample design, and data collection and data processing procedures. The manual provides guidance for understanding and using data from all components of the base year and first follow-up.

The ELS:2002 base-year to first follow-up dataset has been produced in both public-use and restricted-use versions (see appendix D for a summary of differences between the public and restricted electronic codebooks [ECBs]). The released data files reflect alteration or suppression of some of the original data. Such edits were imposed to minimize the risk of disclosing the identity of responding schools and individuals. Although the primary focus of this manual is the public-release version of the data as issued in ECB format, much of the information supplied is also applicable to the restricted-use ECB.

Chapter 1 addresses three main topics. First, it supplies an overview of the NCES education longitudinal studies program, thus situating ELS:2002 in the context of the earlier NCES high school cohorts studied in the 1970s, 1980s, and 1990s. Second, it introduces ELS:2002 by delineating its principal objectives. Third, it provides an overview of the base-year and first follow-up study designs. In subsequent chapters, additional topics are addressed: instrumentation (chapter 2), sample design and weighting (chapter 3), data collection methods and results (chapter 4), data preparation and processing (chapter 5), and data file contents (chapter 6). Appendixes provide additional information, including an introduction to the public-use ECB (appendix A), base-year and first follow-up questionnaires (appendix B), documentation for imputed variables (appendix C), information on variables not included in public-use files but available in restricted-use files for licensed users (appendix D), a glossary of terms (appendix E), student questionnaire critical items (appendix F), base-year and first follow-up ECBs (appendix G), cross-cohort comparisons (appendix H), standard errors and design effects (appendix I), and a synopsis of the ELS:2002 first follow-up field test (appendix J).

[1] RTI International is a trade name of Research Triangle Institute.

1.2 Historical Background

1.2.1 NCES Education High School Longitudinal Studies Program

In response to its mandate to "collect and disseminate statistics and other data related to education in the United States" and the need for policy-relevant, nationally representative longitudinal samples of elementary and secondary students, NCES instituted the National Education Longitudinal Studies program. The aim of this continuing program is to study the educational, vocational, and personal development of students at various stages in their educational careers and the personal, familial, social, institutional, and cultural factors that may affect that development.

NCES (and ELS:2002) is authorized by section 406(b) of the General Education Provision Act (20 U.S.C. 1221e) as amended by the Education Sciences Reform Act of 2002. The Education Sciences Reform Act of 2002 replaced the former Office of Educational Research and Improvement (OERI) with the Institute of Education Sciences (IES), in which NCES is now housed.

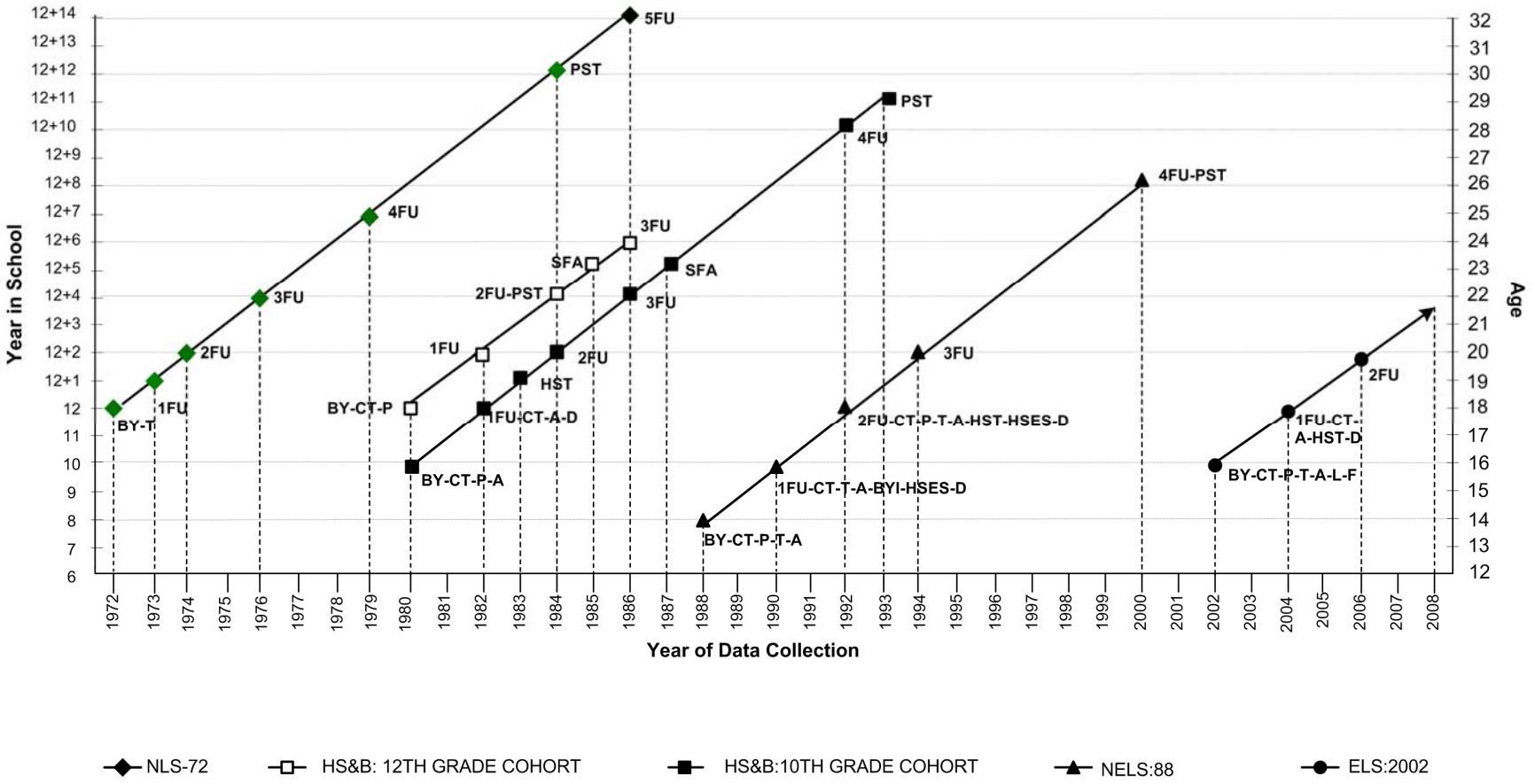
The high school longitudinal studies program consists of three completed studies: the National Longitudinal Study of the High School Class of 1972 (NLS-72), the High School and Beyond (HS&B) longitudinal study of 1980, and the National Education Longitudinal Study of 1988 (NELS:88). In addition, base-year and first follow-up data for ELS:2002, the fourth longitudinal study in the series, are now available. Taken together, these studies describe (or will describe) the educational experiences of students from four decades—the 1970s, 1980s, 1990s, and 2000s—and also provide bases for further understanding of the correlates of educational success in the United States. Figure 1 includes a temporal presentation of these four longitudinal education studies and highlights their component and comparison points. Figure 1 does not identify all future follow-up points for ELS:2002; final decisions have yet to be made concerning them. However, the general expectation is that ELS:2002 sophomores will be followed until about age 30.

1.2.2 National Longitudinal Study of the High School Class of 1972 (NLS-72)

The National Education Longitudinal Studies program began over 30 years ago with the implementation of NLS-72.[2] NLS-72 was designed to provide longitudinal data for education policymakers and researchers who link educational experiences in high school with important downstream outcomes such as labor market experiences and postsecondary education enrollment and attainment. With a national probability sample of 19,001 high school seniors from 1,061 public and religious and other private schools, the NLS-72 sample was representative of approximately 3 million high school seniors enrolled in 17,000 U.S. high schools during the spring of the 1971–72 school year. Each member of this cohort was asked to complete a student questionnaire and a cognitive test battery. In addition, administrators at the sample members'

[2] For documentation on NLS-72, see Riccobono et al. (1981) and Tourangeau et al. (1987). While recent NCES reports and user documentation may be found on the NCES website (http://nces.ed.gov), some older documentation may be unavailable. NLS-72 and older HS&B manuals may be downloaded from the International Archive of Education Data (IAED) at the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan (http://www.icpsr.umich.edu). Materials may also be obtained in microfiche or photocopy format from the Education Resources Information Center (ERIC) database (http://www.eric.ed.gov).

Figure 1. Longitudinal design for the NCES high school cohorts: 2004



NLS-72=National Longitudinal Study of the High School Class of 1972
HS&B=High School and Beyond: 1980
NELS:88=National Education Longitudinal Study of 1988
ELS:2002=Education Longitudinal Study of 2002

BY=Base Year data collection
1FU=1st follow-up data collection
2FU=2nd follow-up data collection
3FU=3rd follow-up data collection
4FU=4th follow-up data collection
5FU=5th follow-up data collection

CT=Cognitive test
P=Parent survey
T=Teacher survey
A=Adminstrator survey
L=Library/media center survey
F=Facilities checklist

HST=High School Transcript
PST=Post-Secondary Transcript
SFA=Student Financial Aid
BYI=Base Year Ineligible Study
HSES=HS Effectiveness Study
D=Dropout Survey

schools were asked to supply information about the schools' programs, resources, and grading systems, as well as survey data on each student. No parent survey was conducted. However, postsecondary education transcripts were collected from the institutions attended by students. Five follow-up surveys were completed with this student cohort, with the final data collection taking place in 1986, when the sample members were 14 years removed from scheduled high school graduation and approximately 32 years old.

A wide variety of data was collected in the NLS-72 surveys. For example, in addition to background information about the students and their families, the base-year and follow-up surveys collected data on each respondent's educational activities (e.g., schools attended, grades received, and degree of satisfaction with educational institutions). Participants were also asked about their work experiences, periods of unemployment, job satisfaction, military service, marital status, and children. Attitudinal information on self-concept, goals, community involvement, and personal evaluations of educational activities were also included in the study.

1.2.3 High School and Beyond (HS&B)

The second in the series of NCES longitudinal studies was launched in 1980. HS&B included one cohort of high school seniors comparable to the NLS-72 sample; however, the study also extended the age span and analytical range of NCES longitudinal studies by surveying a sample of high school sophomores. Base-year data collection took place in the spring term of the 1979–80 academic year with a two-stage probability sample. More than 1,000 schools served as the first-stage units, and 58,000 students within these schools were the second-stage units. Both cohorts of HS&B participants were resurveyed in 1982, 1984, and 1986; the sophomore group also was surveyed in 1992.[3] In addition, to better understand the school and home contexts for the sample members, data were collected from teachers (a teacher comment form in the base year asked for teacher perceptions of HS&B sample members), principals, and a subsample of parents. High school transcripts were collected for a subsample of sophomore cohort members. As in NLS-72, postsecondary transcripts were collected for both HS&B cohorts; however, the sophomore cohort transcripts cover a much longer time span (to 1993).

With the study design expanded to include a sophomore cohort, HS&B provided critical data on the relationships between early high school experiences and students' subsequent educational experiences in high school. For the first time, national data were available that showed students' academic growth over time and how family, community, school, and classroom factors promoted or inhibited student learning. Researchers were able to use data from the extensive battery of achievement tests within the longitudinal study to assess growth in knowledge and cognitive skills over time. Moreover, data were then available to analyze the school experiences of students who later dropped out of high school and, eventually, to investigate their later educational and occupational outcomes. These data became a rich resource

[3] For a summation of the HS&B sophomore cohort study, see Zahs et al. (1995). For further information on HS&B, see the NCES website: http://nces.ed.gov/surveys/hsb/.

for policymakers and researchers over the next decade and provided an empirical base to inform the debates of the education reform movement that began in the early 1980s.[4]

1.2.4 National Education Longitudinal Study of 1988 (NELS:88)

Much as NLS-72 captured a high school cohort of the 1970s and HS&B captured high school cohorts of the 1980s, NELS:88 was designed to study high school students of the 1990s—but with a premeasure of their achievement and status, prior to their entry into high school. NELS:88 represents an integrated system of data that tracked students from junior high or middle school through secondary and postsecondary education, labor market experiences, and marriage and family formation. Because ELS:2002 repeats so many of its innovations and design features, it will be useful to provide a detailed, round-by-round picture of NELS:88.

Data collection for NELS:88 was initiated with the 8th-grade class of 1988 in the spring term of the 1987–88 school year. Along with a student survey, NELS:88 included surveys of parents (base year and second follow-up), teachers (base year, first follow-up, and second follow-ups), and school administrators (base year, first follow-up, and second follow-up). The sample was also surveyed after scheduled high school graduation, in 1994 and 2000.[5]

1.2.4.1 NELS:88 Base Year

The NELS:88 base year (1988) successfully surveyed 24,599 students, out of some 26,432 selected 8th-graders, across 1,052 public, Catholic, and other private schools. In addition to filling out a questionnaire, students also completed assessments in four subjects (mathematics, science, reading, and social studies). The base year also surveyed one parent, two teachers, and the principal of each selected student. The base-year research instruments collected information about home, school, and individual factors that could serve as predictors for later outcomes (such as, viewed in terms of positive outcomes, graduating from high school, making a smooth transition into the workforce, or completing postsecondary education). Information collected in the base year included family income, parental education, and occupation; parental aspirations for their 8th-grader; the 8th-grader's educational and occupational aspirations and plans, school experiences, extracurricular activities, jobs and chores, television viewing, and reading; teacher perceptions of the 8th-grader's classroom performance and personal characteristics; curricular and instructional information about the classes in which teachers taught the 8th-grader; the teacher's own background and activities; and the principal's reports on the educational setting and environment of the school.

[4] For a summary of reforms instituted between the time the HS&B cohort was in high school and the NELS:88 cohort was in middle/junior high and high school, see Rasinski et al. (1993). For a summary of state education reforms instituted during the earlier school years of the ELS:2002 cohort, see Hurst et al. (2003).

[5] The entire compass of NELS:88, from its baseline through its final follow-up in 2000, is described in Curtin et al. (2002). More detailed information about the high school surveys of NELS:88 can be found in Ingels et al. (1994). Final outcomes for NELS:88 (in 2000) are reported in Ingels et al. (2002). The most extensive documentation of the NELS:88 assessment battery is found in Rock and Pollack (1995a). The quality of NELS:88 data in the in-school rounds is examined in Kaufman and Rasinski (1991) and McLaughlin and Cohen (1997). The sample design is documented in Spencer et al. (1990). Eligibility and exclusion issues are addressed in Ingels (1996). NCES keeps an updated version of the NELS:88 bibliography on its website. The bibliography encompasses both project documentation and research articles, monographs, dissertations, and paper presentations employing NELS:88 data (see http://nces.ed.gov/surveys/nels88/Bibliography.asp).

1.2.4.2 NELS:88 First Follow-up

A first follow-up took place in 1990. In the NELS:88 first follow-up (initial data release), there are 19,260 participants (18,220 students and 1,040 dropouts) from a sample of 20,700. (There were some changes to the file in the second follow-up re-release of the 1990 data, which shows a revised sample size of 20,840). At that time, student cohort members, their teachers, and their principals were resurveyed. The first follow-up presented three major new analytic opportunities: (1) longitudinal analysis of gains in tested achievement and the correlates of achievement gains, (2) identification of high school dropouts and investigation of why some students drop out of school and others persist, and (3) cross-cohort comparison (1990 high school sophomores could be compared to sophomores in 1980).

Achievement gain. One major goal of NELS:88 was to measure students' academic growth over time and to identify the specific school (and nonschool) processes that may foster academic achievement. The first follow-up tests were tailored to students' ability as measured in the base year; more difficult test forms were assigned to students with a higher ability estimate. The first follow-up, by retesting the 8th-grade NELS:88 cohort, was able to measure cognitive gains between 8th and 10th grades in mathematics, science, reading, and social studies. In turn, these gains could be related to the data collected on home and school correlates of achievement, starting in 1988. Because NELS:88 developed hierarchical criterion-referenced proficiency scores (in reading, science, and mathematics), gain can be looked at in more than just quantitative terms—one can use the proficiency levels to locate the place on the growth continuum where the gain took place (e.g., at a lower or at a higher skill area) and, in turn, better relate gains to specific school processes and curricular sequences.[6]

Dynamics of school disengagement and dropping out. Another major goal of the first follow-up was to study the educational trajectory of those who drop out of high school and to better understand the factors that help some at-risk students persist in their education. By beginning with the 8th grade, NELS:88 was able to capture the population of early dropouts—those who left school prior to spring term of 10th grade—as well as (in the second follow-up) later dropouts (who left after spring of 10th grade) as had been studied in HS&B.

Cross-cohort comparison. A third goal of the 1990 wave was to compare NELS:88 sophomores with the earlier cohort of high school sophomores studied in HS&B. To ensure comparability of the two samples, NELS:88 "freshened" the sophomore sample by giving a chance of selection to 1990 sophomores who had not been 8th-graders in 1988 (or had not been in the United States). Thus, a nationally representative sophomore cohort was included in NELS:88 in the first follow-up (1990).

1.2.4.3 NELS:88 Second Follow-up

The second follow-up took place in the spring term of the 1991–92 school year, when most sample members were in their final semester of high school. There were 21,188 student and dropout participants. This follow-up provided a culminating measurement of learning in the course of secondary school and also collected information to facilitate investigation of the

[6] Further information about NELS:88 proficiency scores can be found in Rock and Pollack (1995a). For examples of their use in achievement gain analysis, see Rock and Pollack (1995b) and Scott et al. (1995).

transition into the labor force and postsecondary education after high school. As in the first follow-up, the sample was freshened, this time to represent the high school senior class of 1992. Trend comparisons can be made to the high school classes of 1972 and 1980 that were studied in NLS-72 and HS&B. The NELS:88 second follow-up also surveyed students who were identified as dropouts in 1990 and identified and surveyed additional students who had left school since the prior wave. In late 1992 and early 1993, high school transcripts were collected for sample members.

1.2.4.4 NELS:88 Third Follow-up

The third follow-up took place in 1994, when most sample members had completed high school. The primary goals of the 1994 round were (1) to provide data for trend comparisons with NLS-72 and HS&B, (2) to address issues of employment, (3) to address issues of postsecondary access and choice, and (4) to ascertain how many dropouts had returned to school and by what route. There were 14,915 participants.

1.2.4.5 NELS:88 Fourth Follow-up

The fourth follow-up took place in 2000, when most sample members who attended college and technical schools had completed their postsecondary education. The study data address issues of employment, family formation, and postsecondary persistence and attainment. There were 12,144 participants in the questionnaire phase of the study. In fall 2000 and early 2001, postsecondary transcripts were collected, further increasing the analytic potential of the data and the possibility of examining trends over time.

1.3 Education Longitudinal Study of 2002 (ELS:2002)

ELS:2002 represents a major longitudinal effort designed to provide trend data about critical transitions experienced by students as they proceed through high school and into postsecondary education or their careers. The 2002 sophomore cohort is being followed, initially at 2-year intervals, to collect policy-relevant data about educational processes and outcomes, especially as such data pertain to student learning, predictors of dropping out, and high school effects on students' access to, and success in, postsecondary education and the workforce.

In the spring term 2002 base year of the study, high school sophomores were surveyed and assessed in a national sample of high schools with 10th grades. Their parents, teachers, principals, and librarians were surveyed as well.

In the first of the follow-ups, base-year students who remained in their base-year schools were resurveyed and tested (in mathematics) 2 years later, along with a freshening sample that makes the study representative of spring 2004 high school seniors nationwide. Students who had transferred to a different school, had switched to a homeschool environment, graduated early, or who had dropped out were administered a questionnaire.

This section introduces ELS:2002, lists some of the major research and policy issues that the study addresses, and explains the four levels of analysis—cross-sectional, longitudinal, cross-cohort, and international comparison—that can be conducted with ELS:2002 data.

1.3.1 ELS:2002 Study Objectives

ELS:2002 is designed to monitor the transition of a national sample of young people as they progress from 10th grade through high school and on to postsecondary education or the world of work, or both.

ELS:2002 has two distinctive features. First, it is a longitudinal study, in which the same units (schools and students) are surveyed repeatedly over time. Individual students will be followed through high school and for a number of years thereafter. The base-year schools were surveyed twice, in 2002 and in 2004. Second, in the high school years, ELS:2002 is an integrated, multilevel study that involves multiple respondent populations. The respondents include students, their parents, their teachers, and their schools (from which data are collected at four levels: from the principal, the librarian, a facilities checklist, and school course catalogues and records, which will support a course offerings component in the first follow-up transcript study). Each of the two distinctive features—the longitudinal nature of the ELS:2002 design and its multilevel focus—will be explained in greater detail below.

The transition through high school and beyond into postsecondary institutions and the labor market is both complex (youth may follow many different paths) and prolonged (it takes place over a period of years). The complexity and time frame for this transition make longitudinal approaches especially appropriate. By surveying the same young people over time, it is possible to record the changes taking place in their lives. It is also possible to gather information about the ways that earlier achievements, aspirations, and experience predict what happens to the respondents later. In the baseline data collection (spring 2002), ELS:2002 measured students' tested achievement in reading and mathematics. ELS:2002 also obtained information from students about their attitudes and experiences.

These same students have been resurveyed 2 years later (in 2004), in the ELS:2002 first follow-up, to measure changes such as achievement gains in mathematics and changes in enrollment status (e.g., the situation of students who drop out of school compared with those who persist in their education).

Cohort members will be followed for a number of years thereafter so that later outcomes (e.g., their access to and persistence in higher education or their success in the labor market) can be understood in terms of their earlier aspirations, achievement, and high school situation.

ELS:2002 gathers information at multiple levels. It obtains information not only from students and their school records, but also from students' parents, teachers, and the administrators (principal and library media center director) of their schools. Data from their teachers, for example, provide information both about the student's and the teacher's backgrounds and activities. This multilevel focus supplies researchers with a comprehensive picture of the home, community, and school environments and their influences on the student.

This multiple-respondent perspective is unified by the fact that, for most purposes, the student is the basic unit of analysis.[7]

Using this multilevel and longitudinal information, the base year (2002) and first follow-up (2004) of ELS:2002 will help researchers and policymakers explore and better understand such issues as the importance of home background and parental aspirations for a child's success; the influence of different curriculum paths and special programs; the effectiveness of different high schools; and whether a school's effectiveness varies with its size, organization, climate or ethos, curriculum, academic press, or other characteristics. These data will facilitate understanding of the impact of various instructional methods and curriculum content and exposure in bringing about educational growth and achievement.

After the high school years, ELS:2002 will continue to follow its sample of students into postsecondary education or the labor market, or both. For students who continue on to higher education, researchers can use ELS:2002 to measure the effects of their high school careers on subsequent access to postsecondary institutions, their choices of institutions and programs, and as time goes on, their postsecondary persistence, attainment, and eventual entry into the labor force and adult roles. For students who go directly into the workforce (whether as dropouts or high school graduates), ELS:2002 will be able to determine how well high schools have prepared these students for the labor market and how they fare within it.

Key elements in the ELS:2002 longitudinal design are summarized by wave below.

Base Year (2002)

- Completed baseline survey of high school sophomores in spring term 2002.
- Completed cognitive tests in reading and mathematics.
- Completed survey of parents, English teachers, and mathematics teachers. Collected school administrator questionnaires.
- Included additional components for this study—a school facilities checklist and a media center (library) questionnaire.
- Established sample sizes of approximately 750 schools and over 17,000 students. Schools are the first-stage unit of selection, with sophomores randomly selected within schools.
- Oversampled Asian[8] and Hispanic students and private schools.
- Designed linkages with the Program for International Student Assessment (PISA); scored reporting linkages to the prior longitudinal studies.

[7] Base-year school administrator, library media center, and facilities data can be used to report on the nation's schools with 10th grades in the 2001–02 school year. A first follow-up course offerings file will further enrich the information available about high schools with 10th grades in 2002. However, if history is a guide, most analysts will employ the school-level data to provide further contextual information on the student.

[8] Except where indicated otherwise, the race/ethnicity variable for this report includes six categories: (1) American Indian or Alaska Native; (2) Asian or Pacific Islander, including Native Hawaiian; (3) Black, including African American; (4) Hispanic or Latino; (5) More than one race; and (6) White. All race categories exclude individuals of Hispanic or Latino origin.

First Follow-up (2004)

- Most sample members were seniors, but some were dropouts or in other grades (early graduates or retained in an earlier grade).
- Student questionnaire (different versions for students who remained in the base-year school, transferred to a new school, completed high school early, or were homeschooled), dropout questionnaire, assessment in mathematics, and school administrator questionnaire were administered.
- Returned to the same schools but separately followed transfer students and surveyed them outside of school.
- Freshened for a senior cohort.
- High school transcript component in 2004 (coursetaking records at the student level for grades 9–12) and course offerings component at the school level.

Second Follow-up (2006)

- Post-high-school follow-up with web-based instrument for self-administration, computer-assisted telephone interview (CATI), or computer-assisted personal interview (CAPI).
- Survey 2 years after scheduled high school graduation.

Further Follow-ups

- Number of (and dates for) further web/CATI/CAPI follow-ups to be determined.

1.3.2 ELS:2002 Research and Policy Issues

Apart from helping to describe the status of high school students and their schools, ELS:2002 will provide information to help address a number of key policy and research questions. The study is intended to produce a comprehensive dataset for the development and evaluation of education policy at all government levels. Part of its aim is to inform decisionmakers, education practitioners, and parents about the changes in the operation of the educational system over time and the effects of various elements of the system on the lives of the individuals who pass through it. Issues that can be addressed with data collected in the high school years include the following:

- students' academic growth in mathematics;
- the process of dropping out of high school;
- the role of family background and the home education support system in fostering students' educational success;
- the features of effective schools;
- the relationship between coursetaking choices and success in the high school years (and thereafter);

- the distribution of educational opportunities as registered in the distinctive school experiences and performance of students from various subgroups. Such subgroups include the following:
  - students in public and private high schools;
  - language minority students;
  - students with disabilities;
  - students in urban, suburban, and rural settings;
  - students in different regions of the country;
  - students from upper, middle, and lower socioeconomic status levels;
  - male and female high school students; and
  - students from different racial or ethnic groups.
- steps taken to facilitate the transition from high school to postsecondary education or the world of work.

After ELS:2002 students have completed high school, a new set of issues can be examined. These issues include the following:

- the later educational and labor market activities of high school dropouts;
- the transition of those who do not go directly on to postsecondary education or to the world of work;
- access to and choice of undergraduate and graduate educational institutions;
- persistence in attaining postsecondary educational goals;
- rate of progress through the postsecondary curriculum;
- degree attainment;
- barriers to persistence and attainment;
- entry of new postsecondary graduates into the workforce;
- social and economic rate of return on education to both the individual and society; and
- adult roles, such as family formation and civic participation.

These research and policy issues can be investigated at several distinct levels of analysis. The overall scope and design of the study provide for the four following analytical levels:

- cross-sectional profiles of the nation's high school sophomores (2002), seniors (2004), and post-sophomore-year dropouts (2004);
- longitudinal analysis (including examination of life course changes);
- intercohort comparisons with American high school students of earlier decades; and

- international comparisons: U.S. 15-year-olds to 15-year-olds in other nations, including longitudinal outcomes for the United States that can be related to scale scores in mathematics and reading from PISA.

1.3.2.1 Cross-Sectional Profiles

Cross-sectional data permit characterization of the nation's high school sophomores in the spring term of the 2001–02 school year. Initial cross-sectional findings from the base year are available in an NCES report, *A Profile of the American High School Sophomore in 2002*.[9] Because of sample freshening, the results 2 years later provide a basis for profiling the nation's high school seniors in the spring term of the 2003–04 school year. A report on seniors is now being prepared for release.

1.3.2.2 Longitudinal Analysis

Longitudinal analysis has become possible with the release of data from the 2004 first follow-up. The primary research objectives of ELS:2002 are longitudinal in nature. The study provides the basis for within-cohort comparison by following the same individuals over time to measure achievement growth in mathematics and monitor enrollment status over the high school years. With future follow-ups, the study can also record such key outcomes as postsecondary entry and attainment, labor market experiences, and family formation. In turn, these outcomes can be related to antecedents identified in earlier rounds, including individual, home, school, and community factors.

1.3.2.3 Intercohort Comparisons

As part of an important historical series of studies that repeats a core of key items each decade, ELS:2002 offers the opportunity for the analysis of trends in areas of fundamental importance, such as patterns of coursetaking, rates of participation in extracurricular activities, academic performance, and changes in goals and aspirations. With completion of the first follow-up in 2004, researchers can now compare ELS:2002 high school seniors' experiences, attitudes, and achievement with that of NELS:88 seniors in 1992, HS&B seniors in 1980, and NLS-72 seniors in 1972. They will also be able to compare ELS:2002 dropouts in 2004 with the high school dropouts studied by HS&B in 1982 and by NELS:88 in 1992.

Upon release of ELS:2002 academic transcript data, trend comparisons can also be made with academic transcript data containing students' high school course histories and sequences, since comparable transcript studies have been conducted, starting with HS&B (1982) and including NELS:88 (1992) and the National Assessment of Educational Progress (NAEP) (1987, 1990, 1994, 1998, and 2000).

1.3.2.4 International Comparisons

A feature of ELS:2002 that expands the study's power beyond that of the predecessor studies is that it will be used to support international comparisons. Items have been included on

[9] See Ingels et al. (2005). A small, but growing, ELS:2002 bibliography can be found at http://nces.ed.gov/surveys/els2002/Bibliography.asp.

the ELS:2002 achievement tests from PISA. The Organization for Economic Cooperation and Development's (OECD's) PISA (Lemke et al. 2001) is an internationally standardized assessment, jointly developed by the 32 participating countries (including the United States) and administered to 15-year-olds in groups in their schools. PISA covers three domains: reading literacy, numeracy, and scientific literacy; ELS:2002 test results have been linked to PISA reading and mathematics scores so that the PISA scale can be used in ELS:2002 analyses. PISA aims to define each domain not merely in terms of mastery of the school curriculum, but also in terms of important knowledge and skills needed in adult life. Emphasis is placed on the mastery of processes, the understanding of concepts, and the ability to function in various situations within each domain.

1.3.3 Overview of the Base-Year Study Design[10]

The ELS:2002 base-year study was carried out in a national probability sample of 752 public, Catholic, and other private schools in the spring term of the 2001–02 school year. Of 17,591 eligible selected sophomores, 15,362 completed a base-year questionnaire, as did 13,488 parents, 7,135 teachers, 743 principals, and 718 librarians.

Seven study components comprise the base-year design: assessments of students (achievement tests in mathematics and reading); a survey of students; surveys of parents, teachers, school administrators, and librarians; and a facilities checklist (completed by survey administrators, based on their observations at the school). The student assessments measured achievement in mathematics and reading; the baseline scores can serve as a covariate or control variable for later analyses. Mathematics achievement was reassessed 2 years hence, so that achievement gain over the last 2 years of high school can be measured and related to school processes and mathematics coursetaking. The student questionnaire gathered information about the student's background, school experiences and activities, plans and goals for the future, employment and out-of-school experiences, language background, and psychological orientation toward learning.

One parent of each participating sophomore was asked to respond to a parent survey. The parent questionnaire was designed to gauge parental aspirations for their child, home background and the home education support system, the child's educational history prior to 10th grade, and parental interactions with and opinions about the student's school. For each student enrolled in English or mathematics, a teacher was also selected to participate in a teacher survey. The teacher questionnaire collected the teacher's evaluations of the student and provided information about the teacher's background and activities. The head librarian or media center director at each school was asked to complete a library media center questionnaire, which inquired into the school's library media center facility, its staffing, its technological resources, collection and expenditures, and scheduling and transactions. Finally, the facilities checklist was a brief observational form completed for each school. The form collected information about the condition of school buildings and facilities.

[10] Although this manual covers the base year as well as first follow-up of ELS:2002, much more detailed information about the base year can be found in Ingels et al. (2004).

1.3.4 Overview of the First Follow-up Study Design

The basis for the sampling frame for the first follow-up was the sample of schools and students studied in the ELS:2002 base year. There were two overlapping but conceptually different target student populations, or populations of inferential interest, for the first follow-up. One population (the ELS:2002 sophomore cohort) consists of those students who were enrolled in the 10th grade in the spring term of 2002. The other population (the ELS:2002 senior cohort) comprises those students who were enrolled in the 12th grade in the spring term of 2004. The former population includes students who dropped out of school between 10th and 12th grades, students who graduated early, students who went from a school setting to a homeschooling setting, and students who fell behind the modal grade progression of their peers (e.g., students who repeated a grade and were 11th-graders in spring 2004). Because of these two target populations and the major analytical subgroups, the full-scale sample encompasses the following types of students in the spring of 2004:

- ELS:2002 base-year student sample members enrolled (in either the 12th grade or some other grade) in the school in which they were originally sampled;
- ELS:2002 base-year sophomores who dropped out of school prior to first follow-up (2004) data collection;
- ELS:2002 base-year student respondents who finished high school early, including those who graduated from high school early, as well as those who did not graduate because they achieved alternative certification (e.g., exam-certified equivalency such as a GED);
- ELS:2002 base-year student respondents who transferred out of the school in which they were originally sampled (including homeschooled students);
- ELS:2002 base-year sample students who were deemed unable to participate directly during the base year owing to severe disability or insufficient command of the English language such that they could not complete a questionnaire; and
- students at the base-year sample school who were enrolled in the 12th grade in spring of 2004 but who were not in 10th grade in the United States during the 2001–02 school year. In spring term 2002, such students may have been out of the country, been enrolled in school in the United States in a grade other than 10th, had an extended illness or injury, been homeschooled, been institutionalized, or temporarily dropped out of school. These students comprised the first follow-up "freshening" sample.

While all groups in the sample as categorized above were eligible to complete a questionnaire, different instruments were tailored to different study populations. The guiding intuition was to provide a core of items that all sample members would respond to, supplemented by items specific to the circumstances of a particular group (such as dropouts, for example, for whom questions about their current school situation would not be relevant). In chapter 2, the various questionnaires—student, abbreviated student, transfer student, early graduate, homeschool, out-of-school (dropout), and new student supplement—are described at length.

For some classifications of the sample, a first follow-up test score in mathematics has either been collected (students still in the base-year school) or imputed (students who have transferred to a new school). For other categories of sample members, such as dropouts, early graduates, and the homeschooled, a test score has neither been collected nor imputed. (It should also be noted that missing base-year test score data have been imputed for base-year nonrespondents who became respondents in the first follow-up.)

For all classifications of sample members, information about student coursetaking (covering all years of high school and including the sequence in which courses were taken and grades earned) will be collected late in 2004 and early 2005 through the high school transcript component of the ELS:2002 first follow-up study.

At the school level, the first follow-up has extended information about base-year schools through administration of a school administrator questionnaire. In addition, information about school course offerings will be collected in the first follow-up transcript study. Finally, further information about participating schools at the time of the first follow-up survey can be obtained (on the restricted file only) by linking (via the NCES identification code [NCESID]) to the Common Core of Data (CCD) or Private School Study (PSS), and, via zip codes, to 2000 Census data. The NCES school district database and its Census data also are accessible on the restricted-use file by means of the standard NCES school ID.

Chapter 2
Instrumentation

2.1 Introduction

The base-year (2002) data collection instruments for the Education Longitudinal Study of 2002 (ELS:2002) consisted of five separate questionnaires (student, parent, teacher, school administrator, and library media center), two achievement tests (assessments in reading and mathematics), and a school observation form (facilities checklist).

The first follow-up (2004) data collection instruments comprised seven questionnaires and an achievement test in mathematics. The first follow-up questionnaires included a student questionnaire, a transfer student questionnaire, a new participant student questionnaire (NPSQ), a homeschool student questionnaire, an early graduate questionnaire, a dropout (not currently in school) questionnaire, and a school administrator questionnaire. A new participant supplement (NPS) (repeating questions from the base year) and an abbreviated version of the student questionnaire were also offered.[11] The base-year and first follow-up questionnaires can be found as portable document format (PDF) files on the NCES ELS:2002 website (http://nces.ed.gov/surveys/els2002/).

2.1.1 Instrument Development Process and Procedures

In general, the development and review process for each questionnaire consisted of the following steps:

1. *Sharing of draft data elements.* Draft elements of the questionnaires were shared with other government agencies, policy groups, and interested parties.
2. *Technical review panel (TRP) review.* The ELS:2002 TRP, a specially appointed, independent group of substantive, methodological, and technical experts, reviewed the questionnaires.
3. *National Center for Education Statistics (NCES) review.* The questionnaires underwent interdivisional review at NCES.
4. *Questionnaire revision.* The survey instruments were revised based on reviewer comments.
5. *Writing of justification.* A justification was written for the data elements, noting issue areas, constructs to be measured within each, and items that would be used to measure each construct.
6. *Office of Management and Budget (OMB) review.* The federal OMB reviewed the instruments.
7. *Questionnaire revision.* The questionnaires were revised based on OMB comments.

[11] In fact, the new participant student questionnaire is simply the new participant supplement and abbreviated first follow-up student questionnaire, joined together to create one booklet, for convenience of administration.

8. *Field testing and revision.* The instruments were field tested and revised based on field test results.

Specific assessment items for the base-year mathematics and reading tests and first follow-up mathematics test were typically not subject to these reviews, but the larger assessment framework and goals and the results (as seen in overall item statistics from the field test) were an integral element within the review process and, in particular, the deliberations of the TRP.

The field testing of school enlistment and data collection and processing procedures, questionnaires, and assessments was an especially important step in the development of the full-scale base-year and first follow-up studies. Field test instruments were evaluated in a number of ways. For the questionnaires, field test analyses included evaluation of item nonresponse, examination of test-retest reliabilities, calculation of scale reliabilities, and examination of correlations between theoretically related measures. For the achievement tests in mathematics and reading, item parameters were estimated for both 10th and 12th grade in the base-year field test. Both classical and Item Response Theory (IRT) techniques were employed to determine the most appropriate items for inclusion in the final (base-year main study) forms of the two tests. Psychometric analyses included various measures of item difficulty and discrimination, investigation of reliability and factor structure, and analysis of differential item functioning. In the first follow-up field test, similar classical and IRT psychometric analyses were conducted but with a slightly different end in terms of final format: adaptiveness was ensured through a two-stage test in the base year, whereas the test designed for the first follow-up main study based assignment of form on the prior round ability estimate (as had been done in the National Education Longitudinal Study of 1988 [NELS:88]). The base-year field test report is available from NCES (Burns et al. 2003). Findings of the first follow-up field test are summarized in appendix J of this report.

2.1.2 Instrument Development Goals and Constraints

Since the primary research objectives of ELS:2002 are longitudinal in nature, the first priority was to select the items that would prove most useful in predicting future outcomes as measured in future survey waves, or that would represent near-term (2004) outcomes predicted by base-year (2002) variables.

The second priority was to obtain needed cross-sectional data, whenever consistent with the longitudinal objectives, particularly data that could be used for intercohort comparison with past studies or linkage to certain current data collection efforts. Wherever possible, all ELS:2002 instruments were designed to provide continuity and consistency with the earlier education longitudinal studies of high school cohorts. Where appropriate, ELS:2002 drew from the National Longitudinal Study of the High School Class of 1972 (NLS-72), the High School and Beyond (HS&B) longitudinal study, and, most particularly, NELS:88. In addition, questionnaire and test items were in some cases drawn from other NCES programs, such as the National Assessment of Educational Progress (NAEP) (especially for the assessments), the Program for International Student Assessment (PISA) (for both assessments and psychological scales related to orientation toward learning), the Schools and Staffing Survey (SASS) (particularly but not exclusively for items related to the library media center questionnaire), or the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K) (from which was borrowed the concept of a

facilities checklist). Continuity with ELS:2002's historical predecessors and with other NCES survey and assessment programs was pursued to ensure a common standard of measurement that would permit comparisons and increase the usefulness of the ELS:2002 data. Apart from the intercohort or cross-study comparisons that can be sustained through use of the questionnaire and transcript data, ELS:2002 provides score linkages with the testing programs of PISA, HS&B, and NELS:88.

Although maintaining trend items to support intercohort comparisons was a major aim of instrument development, there was also a need to provide new items to address new areas of policy concern and to reflect recent advances in theory. For example, in the base year in particular, educational technology items were developed to reflect the fact that computers have become a major factor in learning in recent years. Psychological scales that reflect recent work in self-efficacy theory and related areas were also added.

Another consideration in the development of the ELS:2002 instruments was the need to obtain factual information from the best source among the various respondent populations. (This was especially an issue for the base year, in which both parents and students were surveyed.) In some cases, the decision to use the best source resulted in a longer wait to secure the information (e.g., the 2002 student questionnaire was not used to collect information on courses taken or grades, and the 2004 questionnaire was used for this purpose only minimally; academic transcripts are a more reliable source of this information, and they were collected in 2005, after most students had completed high school). In most cases, information has been collected from one source only. However, in a few instances, a particular datum for which there was more than one acceptable source in terms of data quality was deemed to be of such importance that some redundancy between instruments seemed an acceptable price to pay. For example, whereas parents are the best source of information about highest parental educational attainment, the importance of this item was such that it was asked on both the base-year student and parent questionnaires to increase the number of sample members for whom this information would be available (and was asked again, of new participants, in the first follow-up).[12]

Finally, some changes in the law regarding questionnaire content that could be asked of students in a school setting under conditions of implied consent had to be taken into account. Specifically, the Protection of Pupil Rights Amendment (PPRA) proscribes collection of information in the following seven areas when minor students are required to participate in a survey, unless prior written parental consent has been obtained:

1. political affiliations or beliefs of the student or the student's parent;
2. mental and psychological problems of the student or the student's family;
3. sexual behavior or attitudes;

[12] In their analysis of NELS:88 base-year (8th-grade) student and parent reports of parental education, Kaufman and Rasinski (1991) found that, although the number of precise matches between student and parent report was only moderate, validity coefficients were relatively high (father = 0.82, mother = 0.76). McLaughlin and Cohen (1997), in a reanalysis of NELS:88 parent and student data, found the percent matching to be only 55.8 percent on father's education and 56.5 percent on mother's education. Nevertheless, they report polychoric correlations of 0.87 for father's education and 0.84 for mother's, indicating a high degree of convergence between student and parent reports. Student reports increase in quality with age. For high school seniors in HS&B, Fetters, Stowe, and Owings (1984) show validity coefficients of 0.89 for father's education (compared to the 0.82 recorded for 8th-graders by Kaufman and Rasinski [1991]) and 0.85 for mother's education (versus 0.76 for 8th-graders in NELS:88).

4. illegal, antisocial, self-incriminating, or demeaning behavior;
5. critical appraisals of other individuals with whom respondents have close family relationships;
6. legally recognized privileged or analogous relationships, such as those of lawyers, physicians, and ministers; and
7. income.

In addition, when the PPRA was amended in the *No Child Left Behind Act of 2001*, an eighth area was added to the list:

8. religious practices, affiliations, or beliefs of the student or student's parent.

A number of topic areas covered in prior studies, such as HS&B and NELS:88, were therefore excluded from the ELS:2002 base-year and first follow-up student questionnaires, including all items on use of tobacco, alcohol, and drugs and past and present illegal, sexual, or antisocial behaviors, as well as psychological problems and appraisals of family members. A few additional items retained on the base-year student questionnaire that later raised PPRA concerns were suppressed from the final dataset (this fact accounts for the several gaps in the questionnaire and variable name number sequences for the base-year student survey).

Basic elements that are or will be encompassed in the ELS:2002 research instruments can be classified in three broad categories:

- background information (normally collected in the base year only, except for respondents first entering the sample in a later round);
- process information (information about possible influences on the student in the home, school, and community environment, as he or she moves through secondary school and beyond into the world of postsecondary education and the adult workforce); and
- outcome information (the eventual outcomes of the transition process, including later educational attainment and labor market status). The base-year questionnaires are rich in background and process items. The first follow-up questionnaires inquire both into process and outcomes, while also establishing a new baseline for examining the transition out of high school.[13] The focus of the final waves of the study will be outcome data.

[13] An example of a first follow-up item that did not appear on the base-year student questionnaire and was intended to help chart the transition from high school to postsecondary enrollment is F1S52, which asks about the importance of various factors in choosing a postsecondary institution. An example of a new item designed to measure an outcome would be F1S14, which inquires into academic progress as judged by whether the student has remained in modal sequence and is now in 12th grade. An example of a repeated measure would be the educational expectation and life values questions. Ultimately, these plans and expectations can be related to future educational, occupational, and social outcomes. Since these items were asked in the base year, are re-asked in the first follow-up, and will be re-asked again in future follow-ups, they help provide a basis for examining the stability of values and goals over time. Finally, because most of these items have been used with the prior NCES longitudinal high school cohorts, they provide a basis for comparing the goals and values of sophomores in 2002 and seniors in 2004 with earlier cohorts, including high school seniors in 1972, 1980, 1982, and 1992.

2.2 Base-Year and First Follow-up Questionnaires

2.2.1 Base-Year Questionnaires

2.2.1.1 Student Questionnaire

The ELS:2002 base-year student questionnaire was typically self-administered. Sophomore sample members normally completed the questionnaire in a group administration in their schools. A small number of students were surveyed outside of school, with a shortened version of the questionnaire in a computer-assisted telephone interview (CATI). Assessments in reading and mathematics were given at the same time (i.e., during the group administration), in a two-stage process in which the first stage was a routing test. The full questionnaire was available only in English, although a shortened Spanish version was also produced.

The student questionnaire was divided into seven sections: (1) locating information, (2) school experiences and activities, (3) plans for the future, (4) non-English language use, (5) money and work, (6) family, and (7) beliefs and opinions about self.

2.2.1.2 Base-Year Parent Questionnaire

The parent questionnaire was to be completed by the parent or guardian most familiar with the sophomore's school situation and experience. Guided by this definition of the preferred respondent, the parent survey respondent was self-selected.

The parent questionnaire was available in both English and Spanish. Both a hardcopy version and an electronic CATI version[14] were produced. The parent questionnaire addressed the following five topic areas: (1) family background, (2) their child's school life, (3) their child's family life, (4) their opinions about their child's school, and (5) their aspirations and plans for their child's future.

2.2.1.3 Base-Year Teacher Questionnaire

The teacher questionnaire was to be completed by the English teacher and the mathematics teacher of each ELS:2002 sophomore. The teacher questionnaire was designed to illuminate questions of the quality, equality, and diversity of educational opportunity by obtaining information in two content areas:

- *Teacher evaluations of students.* The teacher's assessment of the student's school-related behavior and academic performance and educational and career plans and goals. Respondents completed this section with respect to the sample members they instructed in a particular subject.

[14] The approach to parent telephone interviews in the ELS:2002 base year differed from that followed in NELS:88. In NELS:88, to minimize the possibility of mode of administration effects, the parent was asked to read along in the hardcopy questionnaire as the questions were read over the telephone. The interview was not computer assisted. In ELS:2002, the decision was made to take advantage of the logical consistency editing and other features of CATI, and considerable effort was made to constrain the hardcopy questionnaire to items and formats compatible with a CATI administration. ELS:2002 parents were not interviewed over the telephone with the hardcopy questionnaire in hand. This difference accounts for some variations between the NELS:88 and ELS:2002 parent survey instruments.

- *Teacher background.* Information about the teacher's background and activities (e.g., academic training, subject areas of instruction, years of teaching experience, and participation in professional growth activities).

2.2.1.4 Base-Year School Administrator Questionnaire

The base-year school administrator questionnaire collected information on the school in six areas: (1) school characteristics, (2) student characteristics, (3) teaching staff characteristics, (4) school policies and programs, (5) technology, and (6) school governance and climate. The school administrator data can be used contextually, as an extension of the student data, when the student is the fundamental unit of analysis. At the same time, the ELS:2002 base-year school sample is nationally representative and can stand alone as a basis for generalizing to the nation's regular high schools with sophomores in the 2001–02 school year.

2.2.1.5 Library Media Center Questionnaire

For the school library media center component, the school librarian, media center director, or school administrator supplied information about library media center size, organization, and staffing; technology resources and electronic services; extent of library and media holdings, including both collections and expenditures; and levels of facility utilization, including scheduling for use by students and teachers. Finally, the questionnaire also supplied information about the library media center's use in supporting the school's curriculum, that is, how library media center staff collaborate with and support teachers to help them plan and deliver instruction. Information in the library media center questionnaire can be used as contextual data with the student as the unit of analysis or to generalize to libraries within all regular high schools with 10th grades in the United States in the 2001–02 school year.

2.2.1.6 School Facilities Checklist

Instrumentation for the facilities component comprised a checklist to be completed by the survey administrator. The survey administrator was asked to observe a number of conditions at the school, including the condition of the hallways, main entrance, lavatories, classrooms, parking lots, and surrounding neighborhood. Of special interest were indicators of security (metal detectors, fire alarms, exterior lights, fencing, security cameras, etc.) and maintenance and order (trash, graffiti, clean walls and floors, noise level, degree of loitering, etc.). Information gathered in the facilities checklist can be used as contextual data with the student as the unit of analysis, or data can be used at the school level to generalize to all regular high schools with 10th grades in the United States in the 2001–02 school year.

2.2.2 First Follow-up Questionnaires

2.2.2.1 Introduction

Although assessments will be discussed separately in section 2.3, it is useful, as a point of entry into the first follow-up instrumentation, to consider the fact of test availability in conjunction with the main sample populations for which questionnaires were designed. As table 1 makes clear, not all groups were tested in the first follow-up, nor were test scores imputed for all groups.

Table 1. Assessment availability status, by sample group: 2004

Group (status in 2004)	Base year	First follow-up
2002 sophomores in core (base-year) schools in 2004	Tested[1]	Tested[2]
2002 sophomores in transfer schools in 2004	Tested[1]	Imputed
2004 freshened seniors	Unavailable	Tested[2]
2002 sophomores: 2004 dropouts	Tested[1]	Unavailable
2002 sophomores: 2004 early graduates	Tested[1]	Unavailable
2002 sophomores: homeschooled in 2004	Tested[1]	Imputed

[1] Imputed for base-year nonrespondents.

[2] Imputed for first follow-up participant test noncompleters.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year, 2002" and "First Follow-up, 2004."

The following questionnaires were employed in the ELS:2002 first follow-up: student questionnaire, dropout questionnaire, early graduate questionnaire, transfer student questionnaire, homeschool student questionnaire, and NPSQ. A school administrator questionnaire was also offered. For the ELS:2002 data user, it is necessary to specify which items are common to various questionnaires and which are unique, and how each questionnaire group relates to the analytic populations of interest.

For example, one important analysis population is high school seniors in 2004. These individuals will have completed student, new participant, or transfer student questionnaires, but not dropout or early graduate questionnaires. Even within any of these three questionnaire groupings, only a subset of students will in fact be members of the senior cohort. Another possible population for analysis (particularly for the examination of school effects) consists of students who remained in their base-year school between 2002 and 2004. (In examining this group, one may also wish to take into account the movers who transferred to a new school.)

A further important analytic population is the broader sophomore cohort panel. The base-year to first follow-up combined data file provides a basis for examining the 2002 high school sophomore cohort 2 years later. In determining 2004 outcomes for 2002 sophomores, a comprehensive picture of the cohort's situation can be obtained only by looking at the range of situations represented in the first follow-up questionnaires. Some members of the cohort will have remained in the base-year school and will have completed the student questionnaire. Others will have transferred to a new school and will have completed a transfer student questionnaire. Some will be in a homeschool situation, whereas others may be dropouts or early graduates. One particular subset of the completers of the NPSQ, 12th-grade freshened students, would need to be excluded from investigations of sophomore cohort outcomes in 2004. Although analysis populations are properly selected through flags or universe variables, it is also important for the analyst to know which data elements are shared in common across various questionnaire completion groups. Table 2 provides a crosswalk that shows shared and unique items across the first follow-up questionnaires. (Note that the NPSQ is not included on the crosswalk because its coverage is equivalent to that of the NPS plus the abbreviated student questionnaire.)

Table 2. Crosswalk: First follow-up questionnaire type, by shared and nonshared items: 2004

Question	Student questionnaire	Abbreviated student questionnaire	Transfer	Home-school	Early graduate	Drop-out	New participant supplement	Base-year student or parent questionnaire
What grade are you in	14	14	18	14				
What diploma/certificate	15	15	19	15	19			
Science coursework	16	16	20	16	29	27		
Math coursework	17	17	21	17	30	28		
Confidence in math	18							
Calculators/computers in math	19							
Computer use in math classes	20							
College entrance tests	21	18	22	18				
How studied for college tests	22							
Talent Search	23							
Years participated in Talent Search, etc.	24							
Victimization	25							
Extracurricular activities	26	19	23	20	31			
Time spent on extracurriculars	27	20	24	21	32			
Does school have library	28	21	25					
How often uses school library	29	22	26					
How often uses public library	30	23	27	22	33	49		
Hours on homework	31	24	28	23				
Hours on math homework	32							
Additional reading	33	25	29	24	34	50		
Hours watching TV	34	26	30	25	35	51		
Hours playing video games	35	27	31	26	36	52		
Computer use for schoolwork/other	36	28	32	27	37	53		
Computer use at various locations	37	29	33	28	38	54		
Computer use for fun, school, learn things	38							

See notes at end of table.

Table 2. Crosswalk: First follow-up questionnaire type, by shared and nonshared items: 2004—Continued

Question	Student questionnaire	Abbreviated student questionnaire	Transfer	Home-school	Early graduate	Drop-out	New participant supplement	Base-year student or parent questionnaire
Activities outside of school	39	30	34	29	39	55		
Life values	40	31	35	30	40	56		
How will spend summer	41							
How far in school thinks will get	42	32	36	31	41	57		
How far mother and father wants to go	43	33	37	32	42	58		
Most important thing right after high school	44	34	38	33				
Plan to go to school right after high school	45	35	39	34				
Reasons decided not to go right after high school	46	36	40	35				
Plan to continue education in future	47	37	41	36	44			
Where went for info on college entrance	48							
Type of school will most likely attend	49	38	42	37	45			
To how many schools applied	50	39	43	38	46			
Two most likely schools	51	40	44	39	47			
Importance of school characteristics	52	41	45	40	48			
Plan to work right after high school	53	42	46	41				
Full-time job lined up	54	43	47	42				
Who helped select jobs	55							
Job expects after high school	56	44	48	43				
Job expects at age 30	57	45	49	44	56	66		
How much education for job at age 30	58	46	50	45	57	67		
Ever worked for pay	59	47	51	46				
Hours per week	60	48	52	47				
Hours on weekend	61							
Volunteered in past 2 years	62	49	53	48	58	68		

See notes at end of table.

Table 2. Crosswalk: First follow-up questionnaire type, by shared and nonshared items: 2004—Continued

Question	Student questionnaire	Abbreviated student questionnaire	Transfer	Home-school	Early graduate	Drop-out	New participant supplement	Base-year student or parent questionnaire
Types of volunteer organizations	63							
How often discuss with parents	64	50	54	49				
Friends' plans for after high school	65	51	55	50	59	69		
When began going to this school			15					
Reasons for transferring			16					
Agreement w/ statements re school/teachers			17					
Participated in any school-sponsored activities?				19[1]				
Name of school last attended					18	18		
When last attended high school					20	19		
What grade were you in then					21	20		
How earned GED					24	42		
Why decided to complete GED					25	43		
Earned GED in what state					26	44		
When did you graduate/receive equivalency					27	45		
Why decided to graduate/complete early					28			
Enrolled in postsecondary education since leaving high school					43			
How many jobs since high school					49	59		
Current/most recent job					50	60		
When started this job					51	61		
Still have job					52	62		
When did you leave job					53	63		
How much do you earn					54	64		
Hours per week					55	65		

See notes at end of table.

Table 2. Crosswalk: First follow-up questionnaire type, by shared and nonshared items: 2004—Continued

Question	Student questionnaire	Abbreviated student questionnaire	Transfer	Home-school	Early graduate	Drop-out	New participant supplement	Base-year student or parent questionnaire
Did you pass that grade						21		
Before last left, ever leave						22		
When left for first time						23		
When returned to school						24		
Attended high school in 2002–03						25		
How many school days missed						26		
Reasons for leaving school					22	29		
Good decision					23	30		
What people at school did						31		
What parents did						32		
Things that happened in past 2 years						33		
Ever in alternative program						34		
When entered most recent alternative program						35		
Still enrolled						36		
When left						37		
Who referred to alternative program						38		
Services received from alternative program						39		
How many alternative programs participated in						40		
Plan to get a GED?						41		
Currently taking a GED class						46		
Plan to go back to high school/take GED class						47		
When expects to get GED						48		
Date of birth							1	DOBIRTHP
Sex							2	SEX

See notes at end of table.

Table 2. Crosswalk: First follow-up questionnaire type, by shared and nonshared items: 2004—Continued

Question	Student questionnaire	Abbreviated student questionnaire	Transfer	Home-school	Early graduate	Drop-out	New participant supplement	Base-year student or parent questionnaire
Student is Hispanic							3	BYS15
Student's Hispanic subdivision							4	HISPANIC
Race							5	RACE
Student's Asian subdivision							6	ASIAN
English is student's native language							7	STLANG
Student's native language							8	HOMELANG
English skills							9	BYS70a-d
Enrolled in any U.S. school in spring term of 2002							10	
Ever held back a grade							11	BYP46
Grades repeated							12	BYP48a-k
Lives in household at least half of time							13	BYFCOMP
Mother/female guardian's work							14	OCCMOTH
Father/male guardian's work							15	OCCFATH
Parents' education							16	PARED
Family has items in home							17	BYS84a-j

[1] The sole purpose of this item was to reduce the length of the questionnaire for certain respondents by routing them around the set of dependent items. Therefore, this item is not included in the data file.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year, 2002" and "First Follow-up, 2004."

2.2.2.2 Questionnaire Assignment and Content

First follow-up student questionnaire assignment and content. The student questionnaire was administered to sophomore cohort members who had remained in their base-year school as well as to a freshening sample of 12th-graders in those same schools. Students who completed the student questionnaire also were normally eligible for the first follow-up mathematics assessment. Some students were administered an abbreviated version of the questionnaire (these cases are flagged on the data file). The questionnaire was primarily self-administered in in-school survey sessions, and secondarily, for some students, out of school through CATI or occasionally through mail or field interviews.

Some alterations were required to adapt the paper-and-pencil questionnaire to CATI. Generally, the wording of the paper-and-pencil questions was made more conversational for the telephone interview to facilitate the interviewer-respondent interaction. On occasion, adaptations were made to account for the fact that those interviewed by telephone did not have the benefit of seeing the entire question with all of its elements at once. For example, students were asked to report how much coursework they had taken in various subject areas. Respondents who completed the paper-and-pencil form were able to see the full range of mathematics courses listed more or less in the sequence in which they are taught. In this context, it was clear to respondents that "general math" referred to a basic math course as opposed to a catchall category. However, without the visual cues, telephone respondents may have misinterpreted general math to include all math courses. Therefore, for the telephone interview, general math was moved to the end of the list of math courses. Similar adaptations were required for the other telephone-administered questionnaires as well (transfer student, dropout, and so on). Generally, CATI telephone data collection took place subsequent to in-school data collection. Also, there was more ambiguity about the status (dropout, early graduate, transfer, homeschooled, and so on) of sample members interviewed outside the school setting. For this reason, the CATI interview included a series of screening questions to ensure that the proper questionnaire was administered. Such a screener was also used for field cases subject to in-person interview.

The student questionnaire comprised eight content modules. *Part I* of the questionnaire requested contact information in support of the longitudinal design.

Part II covered the student's school experiences and activities. Data generated from this section provide information about extracurricular participation, computer use in English and math, the transition process from sophomore year to upper-level secondary school, and the relationship of curricular programs and coursetaking to educational achievement and persistence. Some of these data may be viewed as outcomes, influenced by factors studied in the base year, and others as predictors of outcomes in future rounds.

Part III, "How You Spend Your Time," inquired about time usage on homework, TV viewing, video and computer games, computers, nonschool reading, library utilization, and other activities. *Part IV* focused on plans and expectations for the future. It included questions that elicited information about students' educational and life goals and values. *Part V*, on education after high school, contained items on postsecondary planning steps and choice criteria. *Part VI* dealt with plans for work after high school. *Part VII* inquired about working for pay, including hours worked per week. Finally, *Part VIII* consisted of items on community, family, and friends.

First follow-up dropout questionnaire assignment and content. Dropouts were defined as sophomore cohort members who were out of school in the spring term of 2004, who had not received a high school diploma or General Educational Development (GED) credentials, and who had missed 4 or more consecutive weeks not due to accident or illness. Students who had a dropout episode but who had been in school for at least 2 weeks at the time of their school's survey day were administered the student questionnaire. The dropout questionnaire was administered in multiple modalities—self-administration, in-person interviewer administration, and over the telephone by means of CATI.

There was considerable overlap between the student and dropout questionnaires. *Part I* collected locating information for longitudinal follow-up. *Part II* contained items on school experiences and activities. Dropouts were asked questions about the school they last attended and their participation in alternative educational programs. In addition, they were asked to supply their specific reasons for leaving school prior to graduation. They were asked as well about plans to get a GED or return to high school. *Part III* covered time use (reading, library patronage, television, videogames, computer use, and so on). *Part IV* asked about plans and expectations for the future. *Part V* provided information to identify the type and amount of work that dropouts were engaged in. It gathered information about students' work status and history, how much they earned, and how many hours they worked. *Part VI* asked about volunteer work or community college and the educational behaviors of friends.

Early graduate questionnaire assignment and content. Early graduates were interviewed outside the school setting, in multiple data collection modalities but most commonly by telephone. Early graduates were defined as sophomore cohort members who had graduated from high school or received a GED on or before March 15, 2004. The approach to early graduates differs somewhat across the several NCES high school cohort studies. In HS&B, the group that was captured was high school completers who finished early (i.e., prior to March 1, 1982). In NELS:88 and ELS:2002, an additional group is included, those who completed by alternative means (e.g., GED) prior to their classmates who were in the modal graduation sequence. In both HS&B and NELS:88, early graduates completed supplementary questions in addition to the full student questionnaire (answering from the vantage point of their recent high school experience). In ELS:2002, early graduates completed only a subset of the items on the student questionnaire, complemented by additional items pertaining to their situation. More specifically, early graduates were asked with whom they consulted when deciding to graduate early, the basis for that decision, and the means by which they did so. They also provided a history of their work and educational experiences since leaving high school.

Transfer student questionnaire assignment and content. Sophomore cohort members who had transferred out of their base-year school to a new school received the transfer student questionnaire. Transfer students were asked a subset of items from the student questionnaire, covering the following topics: school experiences and activities; time use; plans and expectations for the future; education after high school; work after high school; and community, family, and friends. In addition, transfer students were asked when they transferred and their reasons for doing so. Transfer students did not complete a cognitive test, but their test scores have been imputed. Thus, test scores are available for all classes of senior cohort members—sophomore cohort members who were seniors 2 years later regardless of whether they were "movers" or "stayers" and freshened seniors.

Homeschool student questionnaire assignment and content. ELS:2002 does not provide a representative sample of homeschooled high school students. (In the base year, all study sophomores were selected from regular U.S. high schools.) Instead, homeschooled students in ELS:2002 generalize only to sophomores in regular high schools in spring term 2002 who were in a homeschool situation 2 years later. The primary motive for administering a separate questionnaire to this subset of the sophomore cohort was that neither the transfer student questionnaire items nor the dropout items fully fit their situation.

Homeschooled students were asked about their schooling activities and status, including their grade, coursework completed in sciences and math, and steps taken toward college; how they spend their time; their plans and expectations for the future, including education and work after high school; work experiences; and community, family, and friends.

New participant supplement questionnaire (NPSQ) assignment and content; NPS. There are essentially three categories of students who are ELS:2002 new participants in the first follow-up. One class is high school seniors who entered the study through the freshening sample. A second class of new participants is that of base-year nonrespondents who completed a questionnaire in the first follow-up. The third and final class is that of sophomore cohort members who were ineligible in 2002 because of inability to complete a questionnaire, but who were reclassified as capable of completing a questionnaire in 2004. (An example might be an English language learner who was not proficient in English in 2002 but, with 2 additional years of instruction, had reached a level of English proficiency sufficient to deal with the ELS:2002 first follow-up questionnaire.) While the first of these three classes is by definition a student, the second and third groups include both students and out-of-school members of the sophomore cohort (such as dropouts and early graduates).

Any student new to the study at any of the core (base-year) schools was administered the NPSQ. However, transfer students and out-of-school cohort members were administered the relevant questionnaire and a NPS containing the key base-year items. For example, any student new to the study who had transferred to a new school was administered the transfer student questionnaire and a NPS. Any new respondent who was out of school, however, such as a dropout or early graduate, was administered the appropriate out-of-school questionnaire, as well as a NPS. Table 3 summarizes, for all new participants, use of the NPS and NPSQ, as well as base-year and first follow-up assessment status.

The NPSQ gathered information that had been collected (for other students) in the base year on new participants' demographic characteristics, parental education and occupation, and language use. These items are identical to those on the NPS. In addition, a subset of items included on the student questionnaire was also posed to new participants. These items (which are identical in content to the abbreviated student questionnaire) relate to topics such as school experiences and activities; time use; plans and expectations for the future; education and work after high school; and work, community, family, and friendship experiences. In contrast, the NPS gathered the key base-year variables that also were included on the NPSQ.

Table 3. Base-year key variables and test data available, by type of first follow-up new participants: 2004

Group of first follow-up new participants	Source of base-year standard classification variables	Availability of base-year reading and math scores	Availability of first follow-up math scores
Sophomore cohort members in core (base-year) schools in 2004	NPSQ	Imputed	Tested
Sophomore cohort members in new schools in 2004	NPS	Imputed	Imputed
Sophomore cohort members out of school in 2004: dropouts	NPS	Imputed	Unavailable
Sophomore cohort members out of school in 2004: early graduates	NPS	Imputed	Unavailable
Freshened 2004 seniors	NPSQ	Unavailable	Tested
Sophomore cohort members homeschooled in 2004	NPS	Imputed	Unavailable

NOTE: NPSQ = New Participant Supplement Questionnaire. NPS = New Participant Supplement; this instrument contains only the key base-year items.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year, 2002" and "First Follow-up, 2004."

School administrator questionnaire content and content linkages. The school administrator questionnaire collected information on the school in four areas: school characteristics, structure, and policies; student characteristics and programs; teacher and library staff characteristics; and principal reports on the school environment. It should be noted that many school-level variables of analytic interest also pose high risk of disclosure of school identities. For this reason, a number of analysis variables have been limited to the restricted-use file or may be accessed through a link provided only on the restricted-use file.[15]

It should also be noted that school-level data are not nationally representative of American high schools in 2004, since the first follow-up sample did not factor in "births" of new schools and "deaths" of existing schools between 2002 and 2004. First follow-up school data, however, do provide a statistical portrait of a nationally representative sample of American high schools with 10th grades in 2002 (2 years later).

[15] An example of the latter is the link to the NCES Common Core of Data (CCD) and Private School Survey (PSS) provided via the NCES identification code (NCESID). An analyst with a restricted-use license could import into the analysis such variables as, for example, *grade span* (highest grade and lowest grade of school for any of the relevant academic years); *percent minority*; proportion *free lunch qualifiers*; *enrollment*; grade 9 enrollment (2000–01), grade 10 enrollment (2001–02), grade 11 enrollment (2002–03), grade 12 enrollment (2003–04); *metropolitan status* (urbanicity): locale code; *student/teacher ratio; FTEs:* total number full-time classroom teachers; *student enrollment:* overall*; school type* (regular, vocational, special education, other), and so on. A further example of such a restricted-use link is to school zip code, which permits locale variables to be imported from the 2000 decennial Census.

2.3 Base-Year to First Follow-up Cognitive Test Battery

2.3.1 Base-Year Reading and Mathematics Assessments

The purpose of the ELS:2002 assessment battery is to provide measures of student achievement in reading and mathematics that can be related to student background variables and educational processes, for individuals and for population subgroups. The reading and mathematics tests must provide accurate measurement of the status of individuals at a given point in time. The mathematics test must provide accurate measurement of their cognitive growth over time. Assessment data in ELS:2002 will be used to study factors that contribute to individual and subgroup differences in achievement.

2.3.1.1 Test Design and Format

Test specifications for ELS:2002 were adapted from frameworks used for NELS:88. Math tests contained items in arithmetic, algebra, geometry, data/probability, and advanced topics and were divided into process categories of skill/knowledge, understanding/ comprehension, and problem solving. Through inclusion of PISA items, the ELS:2002 math tests placed a somewhat greater emphasis on practical applications and problem solving than did the NELS:88 test forms. Reading tests consisted of reading passages of one paragraph to one page in length, followed by three to six questions based on each passage. The reading passages included literary material as well as topics in the natural and social sciences. Several passages required interpretation of graphs. Questions were categorized as reproduction of detail, comprehension, or inference/evaluation. The test questions were selected from previous assessments: NELS:88, NAEP, and PISA. Most, but not all, were multiple choice. The number of items in each stage of the test is indicated in table 4.

Table 4. Number of items in each ELS:2002 base-year test form for assessing achievement in mathematics and reading, by test form: 2002

Form	Mathematics	Reading
Routing test	15	14
Second stage tests		
Form X (low difficulty)	25	16
Form Y (middle difficulty)	27	17
Form Z (high difficulty)	27	15
Form V (single stage: limited time, mathematics only)	23	†

† Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year, 2002."

2.3.2 First Follow-up Assessment

This section describes the development and format of the ELS:2002 first follow-up mathematics assessment, scoring procedures, score descriptions, and summary statistics. It includes a discussion of links (through equating or concordance) with other studies (NELS:88 and PISA:2000).

The purpose of the ELS:2002 assessments is to provide measures of student achievement in reading and mathematics that can be related to student background variables and educational processes, for individuals and for population subgroups. Reading[16] and mathematics assessments were conducted in the sophomore base year; in the first follow-up, only a mathematics test was administered. Assessment data in ELS:2002 will be used to study factors associated with individual and subgroup differences in achievement. The reading and mathematics tests must provide accurate measurement of the status of individuals at a given point in time and, for mathematics, must provide accurate measurement of their cognitive growth over time.

2.3.2.1 Test Design and Format

As with the base-year test design, the specifications for the ELS:2002 first follow-up math test were adapted from frameworks used for NELS:88. Mathematics tests contained items in arithmetic, algebra, geometry, data/probability, and advanced topics and were divided into process categories of skill/knowledge, understanding/comprehension, and problem solving. However, like the base-year test, the ELS:2002 mathematics tests placed a greater emphasis on practical applications and problem solving than did the NELS:88 tests. The test questions were selected from previous assessments: NELS:88, NAEP, and PISA. Items were field tested[17] one year prior to the 10th- and 12th-grade surveys, and some items were modified based on field test results. Final forms were assembled based on psychometric characteristics and coverage of framework categories. In the base year, about 10 percent of math questions were open ended; all of the 12th-grade mathematics questions were presented in multiple-choice format.

The ELS:2002 assessments were designed to maximize the accuracy of measurement that could be achieved in a limited amount of testing time, while minimizing floor and ceiling effects, by matching sets of test questions to initial estimates of students' achievement. (For definitions of floor effects, ceiling effects, and other technical terms, see the glossary in appendix E.) In the base year, this was accomplished by means of a two-stage test. In 10th grade, all students received a short multiple-choice routing test, scored immediately by survey administrators who then assigned each student to a low, middle, or high difficulty second-stage form, depending on the student's number of correct answers in the routing test. In the 12th-grade administration, students were assigned to an appropriate test form based on their performance in 10th grade. Cut points for the 12th-grade low, middle, and high forms were calculated by pooling information from the field tests for 10th and 12th grades in 2001, the 12th-grade field test in 2003, and the 10th-grade national sample. Item and ability parameters were estimated on a common scale. Growth trajectories for longitudinal participants in the 2001 and 2003 field tests were calculated, and the resulting regression parameters were applied to the 10th-grade national sample. Test forms were designed to match the projected achievement levels of the lowest and highest 25 percent, and the middle 50 percent, of the base-year sample 2 years later. An additional test form with a broad range of item difficulty was assembled for administration to follow-up participants who were new to the sample or who had not received a math score in 10th grade. Additions to and deletions from the base-year sample resulted in 23 percent, 42 percent, and 26 percent of the follow-up sample taking the low, middle, and high difficulty forms, respectively, with the

[16] Please refer to base-year documentation (Ingels et al. 2004, NCES 2004–405) for additional information on the 10th-grade reading test.
[17] For more details about the field tests, see Burns et al. (2003) (NCES 2003–03) and appendix J of this manual.

remaining 10 percent taking the broad-band form. Each of the four test forms contained 32 multiple-choice items.

2.3.2.2 Scoring Procedures

Eleven test records were deleted because tests were incomplete (fewer than 10 items answered) and six more because regular response patterns (e.g., all answers were "A," or "ABCABCABC...," etc.) indicated lack of motivation to answer questions to the best of the student's ability.

The scores used to describe students' performance on the direct cognitive assessment are broad-based measures that report performance as a whole. The scores are based on IRT, which uses patterns of correct, incorrect, and omitted answers to obtain ability estimates that are comparable across different test forms.[18] In estimating a student's ability, IRT also accounts for each test question's difficulty, discriminating ability, and a guessing factor.

IRT has several advantages over raw number-right scoring. By using the overall pattern of right and wrong responses to estimate ability, IRT can compensate for the possibility of a low-ability student guessing several difficult items correctly. If answers on several easy items are wrong, a correct difficult item is assumed, in effect, to have been guessed. Omitted items are also less likely to cause distortion of scores, as long as enough items have been answered right and wrong to establish a consistent pattern. Unlike raw number-right scoring, which necessarily treats omitted items as if they had been answered incorrectly, IRT procedures use the pattern of responses to estimate the probability of correct responses for all test questions. Finally, IRT scoring makes it possible to compare scores obtained from test forms of different difficulty. The common items present in overlapping forms and in overlapping administrations (10th grade and 12th grade) allow test scores to be placed on the same scale.

In the first follow-up survey, IRT procedures were used to estimate longitudinal gains in achievement over time by using common items present in both the 10th- and 12th-grade forms. Items were pooled from both the 10th- and 12th-grade administrations and anchored to the IRT scale of the NELS:88 survey of 1988–92. Item parameters were fixed at NELS:88 values for the items that had been taken from the NELS:88 test battery and to base-year values for non-NELS:88 items. In each case, the fit of the follow-up item response data to the fixed parameters was evaluated, and parameters for common items whose current performance did not fit previous patterns were reestimated, along with non-NELS:88 items new to the follow-up tests.

2.3.2.3 Score Descriptions and Summary Statistics

Several different types of scores that can be used to describe students' performance on the cognitive assessment are described in detail below. IRT-estimated number-right scores measure students' performance on the whole item pool. NELS:88-equated number-right scores estimate how a student would have performed on the 1992 mathematics scale of NELS:88. Standardized scores (T-scores) report students' performance relative to their peers. Quartile scores divide the estimated population distributions for convenience in analyzing relationships of cognitive skills with other variables. NELS:88-equated proficiency probabilities estimate the

[18] For an account of Item Response Theory, see Hambleton (1989) or Hambleton, Swaminathan, and Rogers (1991).

probability that a given student would have demonstrated proficiency for each of the five mathematics levels defined for the NELS:88 survey in 1992.[19]

IRT-estimated number right. These scores are estimates of the number of items students would have answered correctly if they had responded to all of the 85 questions in the mathematics item pool (i.e., all items that appeared on any of the mathematics forms) in the 10th- and 12th-grade administrations combined. The ability estimates and item parameters derived from the IRT calibration can be used to calculate each student's probability of a correct answer for each of the items in the pool. These probabilities are summed to produce the IRT-estimated number-right scores. These scores are not integers because they are sums of probabilities, not counts of right and wrong answers.

It is important to note that the item pool for base-year and first follow-up mathematics forms combined differs from the sophomore-only pool used to report scale scores in the base year. The combined sophomore forms contained a total of 72 items, with 13 additional, harder items added in 12th grade to extend the range of the scale. To place base-year and first follow-up scores on the same scale so that gains over time can be measured, the base-year IRT-estimated number-right scores have been replaced with scores on the new 85-item combined scale. Table 5 shows variable names, descriptions, and summary statistics for the IRT-estimated number-right scores in the new metric that applies to both rounds of the survey. First follow-up statistics are reported both for all first follow-up participants and for the subset of students who were in 12th grade at the time of the survey. The samples include all students with test scores, as well as imputed scores for students who were not tested. (For a discussion of imputation in ELS:2002, see chapter 3 of this manual). The reliability of the test scores is a function of the variance of repeated estimates of the IRT ability parameter (within-variance), compared with the variability of the whole sample. For the combined base-year and first follow-up mathematics tests, the reliability was 0.92. This applies to all scores derived from the IRT estimation, including the standardized and quartile scores.

Table 5. Item Response Theory (IRT)-estimated number-right mathematics scores in 85-item metric: 2004

Variable name	Description	Range	Weighted mean	Weighted standard deviation
BYTXMIR2	Mathematics IRT-estimated number right, base year, reestimated on longitudinal scale	0-85	42.2	14.0
F1TXMIR2	Mathematics IRT-estimated number right, longitudinal scale, all first follow-up participants	0-85	48.3	15.1
F1TXMIR2	Mathematics IRT-estimated number right, longitudinal scale, first follow-up participants who were in 12th grade	0-85	48.6	15.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year, 2002" and "First Follow-up, 2004."

[19] For further information on the NELS:88 proficiency levels, see Rock and Pollack (1995a), *Psychometric Report for the NELS:88 Base Year Through Second Follow-up* (NCES 95–382). For examples of the use of the NELS:88-equated probability proficiency scores in the context of cross-sectional estimation of status in ELS:2002, see chapter 5 of Ingels et al. (2005), *A Profile of the American High School Sophomore in 2002* (NCES 2005–338). For examples of longitudinal use of the probability of proficiency scores (in NELS:88), see chapter 4 of Scott et al. (1995), *Two Years Later: Cognitive Gains and School Transitions of NELS:88 Eighth Graders* (NCES 95–436).

Standardized scores (T-scores). T-scores provide norm-referenced measurements of achievement; that is, estimates of achievement level relative to the population as a whole. A high mean T-score for a particular subgroup indicates that the group's performance is high in comparison with other groups. It does not represent mastery of a particular set of skills, only that the subgroup's mastery level is greater than a comparison group. In other words, T-scores provide information on status compared to students' peers, whereas the IRT-estimated number-right scores represent status with respect to achievement on a particular criterion set of test items. The T-scores can only provide an indicator of the extent to which an individual or a subgroup ranks higher or lower than the national average. The standardized score reported in the database (F1TXMSTD) is a transformation of the IRT theta (ability) estimate, rescaled to a mean of 50 and standard deviation of 10.

Quartile scores. Quartile scores (BYTXMQU) divide the sample into four equal groups, based on the weighted distribution of mathematics scores. Quartile 1 corresponds to the lowest achieving quarter of the population, quartile 4 to the highest.

2.3.2.4 Links to Other Surveys

Scores for ELS:2002 first follow-up are reported on scales that permit comparisons with mathematics data for NELS:88 12th-graders in 1992. In addition, ELS:2002 base-year mathematics scores were linked to the 2003 PISA mathematics scale. (In the base year, ELS:2002 reading scores were put on the PISA [2000] literacy scale; for details see Ingels et al. [2004], NCES 2004–405.) The link to the NELS:88 scales represents a "true" equating. This means that the tests may be considered interchangeable or, in other words, a score on one exam should be equivalent to a score on the other exam. Several conditions must be met for equating two tests. Most important, the tests must measure the same content. Similarity of format, length, reliability, and subgroup performance also supports the interpretation of interchangeable scores.

PISA concordance. The ELS:2002 and PISA mathematics tests did not share enough items to permit common-item equating, so score scales were linked by means of *equipercentile equating* (see Ingels et al. [2004] for a description of the reading equating with PISA). If two exams measuring the same construct are given to two samples from the same population, the score corresponding to a certain percentile on one exam may be considered to be equivalent to the score on the other exam that represents the same percentile of the population. ELS:2002 and PISA test instruments, scoring methods, and populations differed in several respects that affect the equating procedures and interpretation of linked scores.

The most important difference between PISA and ELS:2002 is the definition of the population sampled in each case. Equipercentile equating assumes that the two samples being equated come from the same population. However, important differences exist between PISA and ELS:2002 (see table 6). The PISA population was based on age (students born in 1987), whereas ELS:2002's population was based on grade (high school sophomores). Although the spring term administration dates for PISA and ELS:2002 overlapped, the range of PISA dates was later in the school year, suggesting the possibility of higher scores due to additional weeks or months of schooling.

Table 6. ELS:2002 and Program for International Student Assessment: Spring 2003 (PISA:spring 2003), by sample characteristics: 2002 and 2003

ELS:2002 sample	PISA:spring 2003 sample
10th-graders only	Different grades
Different ages; modal age = 15	Ages 15.25 to 16.25 years
Testing began in January 2002	Testing began in April 2003
14,543 tested	3,983 tested

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002); and Program for International Student Assessment (PISA).

Because of these differences, subsamples of each group were used to compute equivalent percentiles (see table 7). Transformations were computed based on the 10th-graders from each survey who were within a specified range of ages and testing dates. The resulting transformation was then applied to all ELS:2002 students. To make the PISA sample more nearly equivalent to the ELS:2002 sample, only PISA 10th-graders were used in the equating subsample. To make the ELS:2002 sample more nearly equivalent to the PISA sample, only ELS:2002 students between the ages of 15.25 years and 16.25 years (the approximate age range for PISA examinees) were used. ELS:2002 students who were tested before March 15 or after May 31 were deleted from the equating sample. The restricted samples were intended to be representative of 10th-graders between the ages of 15.25 and 16.25 years.

Table 7. ELS:2002 and Program for International Student Assessment: Spring 2003 (PISA:spring 2003) equating sample: 2002 and 2003

ELS:2002 equating sample	PISA:spring 2003 equating sample
10th-graders only	10th-graders only
15.25- to 16.25-year-olds	15.25- to 16.25-year-olds
Exams offered from March 15 to May 31	Exams offered from April 1 to May 31
Equating sample N = 2,743	Equating sample N = 2,400

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002); and Program for International Student Assessment (PISA).

The equipercentile equating was carried out using five-moment smoothing of the weighted frequency distributions. Plots of the equipercentile-equated scores showed extreme deviations in the tails of the distribution from a trend line based on linear approximation. These deviations are probably due to the methodology employed in PISA scoring: the PISA scores are transformations of normally distributed IRT ability estimates, which, if no shrinkage is imposed, tend to have long tails. The ELS:2002 scores, which are sums of probabilities, do not. As a result, the equipercentile conversion becomes distorted in the tails of the distributions. Throughout most of the score range, a quarter point difference in ELS:2002 math scale corresponds to a difference of 2 to 3 points in the PISA metric. But, in the extreme tails of the distribution, a quarter point difference in ELS:2002 math score corresponds to a difference of 5 to 10 points or more in the PISA metric. For this reason, the equipercentile equating was carried out without the data in the top and bottom tails of each distribution. Then the equipercentile transformation was used to link the scores for the middle 90 percent of the students, and the remaining scores were linked based on the linear approximation of the equating transformation. The cutoff points for using equipercentile versus linear transformation were selected such that

the ELS:2002 to PISA link would be monotonic. Table 8 shows the linking methods for implementing PISA:spring 2003 math scales in ELS:2002.

Table 8. Linking methods for implementing Program for International Student Assessment: Spring 2003 (PISA:spring 2003) math scales in ELS:2002: 2002 and 2003

ELS:2002 scale score range	Equating method	Weighted percent of data
12.60–22.05	Linear approximation	10.0
22.06–51.81	Equipercentile transformation	76.0
51.82–68.90	Linear approximation	14.0

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002); and Program for International Student Assessment (PISA).

Data users should keep in mind that the differences between the ELS:2002 and PISA:spring 2003 tests, scoring methods, and populations mean that the link reported here cannot be considered to be a true equating. Although procedures were carried out to compensate for population differences and scoring methods, no claim is made that the scores may be treated as equivalent. It is more appropriate to refer to this link as a concordance: the PISA-scale score represents the score level achieved by students of the same percentile rank in two populations that were matched as closely as possible given the differences described above.

NELS:88-equated scores. Equating the ELS:2002 scale scores to the NELS:88 scale scores was completed through common-item or *anchor equating*. The ELS:2002 and NELS:88 mathematics tests shared 44 math items. These common items provided the link that made it possible to obtain ELS:2002 student ability estimates on the NELS:88 ability scale. (The ELS:2002 data for 12 additional math items did not fit the NELS:88 IRT parameters, so these items were not treated as common items for the purpose of equating.) Parameters for the common items were fixed at their NELS:88 values, resulting in ability estimates consistent with the NELS:88 metric.

The NELS:88-equated IRT-estimated number-right scores for mathematics are estimates of the number of items students would have answered correctly had they taken the NELS:88 exam and responded to all items in the mathematics items pool. The NELS:88 item pool contained 81 mathematics items in all test forms administered in grades 8, 10, and 12. Table 9 shows mathematics scores for ELS:2002 students, reported on the NELS:88 score scale.

Proficiency probability scores. The criterion-referenced NELS:88-equated proficiency probability scores are based on clusters of items that mark different levels on the mathematics scale. Clusters of four items were identified in the NELS:88 tests that marked five hierarchical levels in mathematics:

1. simple arithmetical operations on whole numbers;
2. simple operations with decimals, fractions, powers, and roots;
3. simple problem solving requiring the understanding of low-level mathematical concepts;
4. understanding of intermediate-level mathematical concepts and/or multistep solutions to word problems; and

5. complex multistep word problems and/or advanced mathematical material.

The proficiency levels are hierarchical in the sense that mastery of a higher level typically implies proficiency at lower levels. In NELS:88, students were judged to be proficient if three of the four items in a cluster were answered correctly. The NELS:88-equated proficiency probabilities were computed using IRT-estimated item parameters calibrated in NELS:88. Each proficiency probability represents the probability that a student would pass a given proficiency level defined as above in the NELS:88 sample.

Table 9 shows variable names, descriptions, and summary statistics for the NELS:88-equated proficiency probability scores.

Table 9. ELS:2002 Item Response Theory (IRT) National Education Longitudinal Study of 1988 (NELS:88)-equated estimated number-right score and proficiency probability scores: 2004

Variable name	Description	Range	Weighted mean	Weighted standard deviation
F1NELS2M	Mathematics—NELS-equated estimated number right (1992 scale)	0–81	50.10	14.20
F1TX1MPP	Mathematics—level 1	0–1	.96	.12
F1TX2MPP	Mathematics—level 2	0–1	.78	.37
F1TX3MPP	Mathematics—level 3	0–1	.62	.45
F1TX4MPP	Mathematics—level 4	0–1	.35	.41
F1TX5MPP	Mathematics—level 5	0–1	.04	.14

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "First Follow-up, 2004."

Choosing the appropriate score for analysis. The IRT-estimated number-right, standardized scores (T-scores), proficiency, and quartile scores are all derived from the IRT model and are based on all of the student's responses to the mathematics assessment. That is, the pattern of right and wrong answers, as well as the characteristics of the assessment items themselves, are used to estimate a point on an ability continuum, and this ability estimate, theta, then provides the basis for criterion-referenced and norm-referenced scores. The choice of the most appropriate score for analysis purposes should be driven by the context in which it is to be used.

The *IRT-estimated number-right scores* are overall, criterion-referenced measures of status at a point in time. The criterion is the set of skills defined by the framework and represented by the assessment item pool. These scores are useful in identifying cross-sectional differences among subgroups in overall achievement level. They provide a summary measure of achievement useful for correlational analysis with status variables, such as demographics, school type, or behavioral measures, and may be used in multivariate models as well. These scores may also be used as longitudinal measures of overall growth. However, gains made at different points on the score scale have qualitatively different interpretations. For example, students who made 5-point gains by mastering arithmetical operations are learning very different lessons from those gaining 5 points at the high end of the scale by learning more advanced mathematics. Although the gains in number of scale score points may be the same, the interpretation, and the relationship

with other factors such as coursework, can be expected to be quite different. Comparison of gain in scale score points is most meaningful for groups that started with similar initial status.

The standardized scores (T-scores) are also overall measures of status at a point in time, but they are norm-referenced rather than criterion-referenced. They do not answer the question, "What skills do students have?" but rather, "How do they compare with their peers?" The transformation to a familiar metric with a mean of 50 and standard deviation of 10 facilitates comparisons in standard deviation units. For example, an individual with a T-score of 65 (or a subgroup with a mean of 65) has demonstrated achievement one and one-half standard deviations above the national average for 12th-graders, whereas a score of 45 would correspond to half a standard deviation below the norm. These numbers do not indicate whether students have mastered a particular body of material, but rather what their standing is relative to others.

Quartile scores are convenient normative scores for the user who wishes to focus on analysis of background or process variables separately for students at different achievement levels. For example, one might want to compare the school experiences or educational aspirations of students in the lowest mathematics quartile with those of students in the highest quartile group.

NELS:88-equated estimated number-right and *proficiency probability scores* may be used in a number of ways. Because they are calibrated on the NELS:88 scale, they may be used for cross-sectional comparisons of students' mathematics achievement in 2004 compared with their counterparts in 1992. The NELS:88-equated number-right scores reflect performance on the whole pool of 81 NELS:88 mathematics items, whereas the proficiency probability scores are criterion-referenced scores that target a specific set of skills. The mean of a proficiency probability score aggregated over a subgroup of students is analogous to an estimate of the percentage of students in the subgroup who have displayed mastery of the particular skill. The proficiency probability scores are particularly useful as measures of gain, because they can be used to relate specific treatments (such as selected coursework) to changes that occur at different points along the score scale. For example, two groups may have similar gains in total scale score points, but for one group, gain may take place at an upper skill level, and for another, at a lower skill level. One would expect to see a relationship between gains in probability of proficiency at a particular level and curriculum exposure, such as taking mathematics courses relevant to the skills being mastered.