

Chapter 2: Common Core of Data (CCD)

1. OVERVIEW

The Common Core of Data (CCD) is NCES’s primary database on public elementary and secondary education in the United States. Every year the CCD collects information from the universe of state education agencies (SEAs) on all public elementary and secondary schools and education agencies in the United States. The CCD provides descriptive data about staff and students at the school, school district, and state levels. Information about revenues and expenditures is collected at the school district and state levels. Some of the CCD’s component surveys date back to the 1930s. The integrated CCD was first implemented in the 1986–87 school year.

Purpose

To provide basic statistical information on all children in this country receiving a public education from prekindergarten through grade 12 and information on the funds collected and expended for providing public elementary and secondary education. The specific objectives of the CCD are to (1) provide an official listing of public elementary and secondary schools and education agencies in the nation, which can be used to select samples for other NCES surveys; and (2) provide basic information and descriptive statistics on public elementary and secondary schools and schooling.

Components

There are six components to the CCD: the Public Elementary/Secondary School Universe Survey, Local Education Agency Universe Survey, State Nonfiscal Survey of Public Elementary/Secondary Education, National Public Education Financial Survey (NPEFS), School District Finance Survey, and Teacher Compensation Survey. The CCD surveys consist of data submitted annually to NCES by state education agencies in the 50 states, the District of Columbia, the Bureau of Indian Education (BIE) schools¹, the Department of Defense Dependents Schools, Puerto Rico, and the four outlying areas (American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands).

Public Elementary/Secondary School Universe Survey. This survey collects information on all public elementary and secondary schools in the United States. (In the 2007–08 school year, there were 101,565 operating and 2,264 nonoperating public elementary and secondary schools.) Data include the school’s mailing address, telephone number, operating status, locale (ranging from large city to rural), and type (“regular” or focused on a special area such as vocational education). The survey also collects the student enrollment (membership) for every grade taught in the school; number of students in each of five racial/ethnic groups²; number of students eligible for free-lunch programs; and number of classroom teachers, reported as full-time equivalents (FTEs). In the 1998–99 school year,

SURVEY OF THE UNIVERSE OF ELEMENTARY AND SECONDARY SCHOOLS

CCD collects data through these major components:

- Public Elementary/Secondary School Universe Survey
- Local Education Agency Universe Survey (i.e., school district survey)
- State Nonfiscal Survey of Public Elementary/Secondary Education (i.e., state aggregate nonfiscal survey)
- National Public Education Financial Survey (i.e., state-level financial survey)
- School District Finance Survey
- Teacher Compensation Survey

¹ The BIE assumed administration of these schools from the Bureau of Indian Affairs in 2006.

² Student data have been collected in either five or seven racial/ethnic groups since the 2007–08 school year. However, starting in 2011–12, student data will be collected only in seven racial/ethnic groups.

several variables were added: location address (if different from mailing address); Title I, magnet, and charter school status; number of students eligible for reduced-price lunch programs; number of migrant students enrolled the previous year; and enrollment broken out by race and sex within grade.

Local Education Agency Universe Survey. This survey serves as a directory of basic information on local education agencies (LEAs). (In the 2007–08 school year, there were approximately 18,090 LEAs, including 17,941 operating and 149 nonoperating agencies.) It collects the agency’s mailing address, telephone number, county location, metropolitan status, and type. The survey includes, for the current year, the total number of students enrolled (membership) in prekindergarten through grade 12; number of ungraded students; number of English language learner (ELL) students served in appropriate programs; and number of instructional, support, and administrative staff. It includes, for the previous year, the number of high school graduates, other completers, and grade 7–12 dropouts. Dropout data were first collected in the 1992–93 CCD, reflecting dropouts for the 1991–92 school year. In 2006–07, the CCD collected both the prior- and current-year number of high school graduates, other completers, and grade 7–12 dropouts. Since 2007–08, however, only current-year data on high school completers and dropouts have been collected. Also, since 2007–08, the high school dropout and completion data have been separated from the LEA universe survey data and released as standalone data files.

State Nonfiscal Survey of Public Elementary/Secondary Education. This survey collects information on all students and staff aggregated to the state level, including number of students by grade level; counts of FTE staff by major employment category; and high school completers by race/ethnicity. Since 2007–08, data on student enrollment and staffing are for the current school year. Through school year 2005–06, data on high school completers and dropouts were collected for the previous year. The collection cycle for school year 2006–07 was a transition year when both prior- and current-year data on high school completers and dropouts were collected.

National Public Education Financial Survey (NPEFS). This survey collects detailed finance data at the state level, including average daily attendance, school district revenues by source (local, state, federal), and expenditures by function (instruction, support services, and noninstruction) and object (salaries, supplies, etc.). It also reports capital outlay and debt service expenditures. Revenues and expenditures are audited after the close of the fiscal year and are then submitted to NCES by each state education agency.

The NPEFS underwent a major revision in fiscal year (FY) 1989, acquiring its present name in that year and greatly increasing the number of data items collected. Since that year, additional items have been added to and deleted from the survey. In the FY 89 data collection, NCES also began providing “crosswalk” software to assist states in their reporting and to improve the comparability of data across states. This software converts a state’s existing accounting reports to uniform federal standards, as described in the NCES accounting handbook (National Forum on Education Statistics 2003). The most recent change in the NPEFS is the addition of teacher salary expenditures broken out by program (regular, special education, vocational, and other education program), as well as the addition of textbook expenditures. Data on expenditures from the America Reinvestment and Recovery Act will be collected and reported separately for fiscal years 2009 through 2011.

School District Finance Survey. This survey collects detailed data by school district, including revenues by source, expenditures by function and subfunction, and enrollment. These data are collected by the Governments Division of the U.S. Census Bureau and are released as the Annual Survey of Local Government Finances (F-33). Before FY 95, data were collected from all districts in decennial census years (e.g., 1990) and years ending in 2 and 7, and from a large sample in other years. The F-33 was first conducted in FY 80. Beginning with FY 95, detailed fiscal data on revenues and expenditures have been collected for all school districts providing public education to students in prekindergarten through grade 12. These data can be linked to the nonfiscal data collected in the Local Education Agency Universe Survey. Student counts and amounts of debt at the beginning and end of the fiscal year are also provided. NCES began to substantially support the F-33 in FY 92.

In FY 97, two variables, Payments to Private Schools and Payments to Public Charter Schools, were added. In FY 1998, two variables that describe the nature of school districts and their relation to other surveys and data files were added: AGCHRT and CENFILE. AGCHRT identifies school districts with charter schools, and CENFILE identifies those districts that are available in the Census Bureau’s version of the F-33 school district file. Similar to changes in the NPEFS, teacher salary and textbook exhibit items were added to the F-33 beginning with the FY 04 collection. Special exhibit items are separate data items that are included in, but do not summarize to, other data items. Starting with the FY 05 collection, the data item Federal Revenue—Bilingual Education (B11) was moved from the “federal revenue direct” section to the “federal revenue through the state” section. This change was made as a result of changes in the allocation of

bilingual education funds by the U.S. Department of Education. In the FY 06 collection, four new local revenue items were added: rents and royalties, sale of property, fines and forfeits, and private contributions. Data on expenditures from the America Reinvestment and Recovery Act will be collected and reported separately for fiscal years 2009 through 2011.

Teacher Compensation Survey. This survey collects total compensation, teacher status, and demographic data about individual teachers from multiple states. In 2007, NCES launched the pilot Teacher Compensation Survey (TCS) data collection, with seven states volunteering to provide administrative records for school year (SY) 2005-06. The TCS expanded to 17 states reporting SY 2006-07 and SY 2007-08 data. Twenty-three states are currently participating and up to 35 states will volunteer to participate in the TCS from 2010 to 2013. The TCS file can be merged with the CCD Public Elementary/Secondary School Universe Survey file. Unique ID numbers are used to track teachers within states over time. The data are released as a restricted-use file, available to researchers with an IES data license. The data items on the restricted-use file include: Teacher ID, NCES School ID, FTE, base salary, total salary, employee benefits, years of teaching experience, highest degree earned, race, age, and teacher status codes. Teachers at more than one school will have a record for each school they teach in, and the FTE and salary values are for the teacher at that school only. Summary descriptive statistics are released in public use files. The public use files include teachers' mean base salary, level of education, and mean base salary by varying levels of experience at the school and LEA level.

Periodicity

Annual. Some of the component surveys were initiated during the 1930s. In its integrated form, the CCD was introduced in the 1986–87 school year.

2. USES OF DATA

The CCD collects three categories of information: (1) general descriptive information on schools and school districts, including name, address, phone number, and type of locale; (2) data on students and staff, including demographic characteristics (e.g., race/ethnicity); and (3) fiscal data covering revenues and current expenditures. The datasets within the CCD can be used separately or jointly to provide information on many topics related to education. The ease of linking CCD data with other datasets makes the CCD an even more valuable resource.

Not only is the CCD a source of data that can be used to demonstrate relationships between different school,

district, and state characteristics, it can also provide a historical record of schools or agencies of interest. This information can shed light on how and why education in the United States is changing. The types of schools or districts that have changed the most with respect to a measured characteristic (e.g., proportion of Hispanic students) can be identified, and the reasons for these changes can be independently investigated. Similarly, the impact of state and local education policies and practices can be assessed through an examination of changes in school and district characteristics. For example, districts that have shown substantial improvement in their racial balance or interracial exposure indices can be identified. The policies and practices employed by these districts can then be examined. By identifying the presence of significant changes and where these changes are occurring, CCD data can help policymakers and practitioners better target their efforts and help researchers develop more sharply focused hypotheses for investigating key education issues.

3. KEY CONCEPTS

The concepts described below pertain to the levels of data collection (school, agency, state) and school locale in the CCD. For a comprehensive list of CCD terms and definitions, refer to the glossaries in various CCD annual publications (such as CCD files and documentation, First Look reports, and technical user guides) available on the Internet (<http://nces.ed.gov/ccd/ccd.publications>).

Local Education Agency. An LEA has administrative responsibility for providing instruction or specialized services to one or more elementary or secondary schools. Most LEAs are *regular school districts* that are locally administered and directly responsible for educating children. Others are *supervisory unions* (which provide administrative systems for the smaller regular districts with which they are associated); *regional education service agencies* (which offer research, data processing, special education or vocational program management, and other services to a number of client school districts); *state-operated school districts* (e.g., for the deaf and blind); *federally operated school districts* (e.g., operated by the Bureau of Indian Education); and *other agencies* not meeting the definitions of the preceding categories (e.g., operated by a Department of Corrections). Since school year 2007–08, a charter agency type code has been used to differentiate charter agencies from other types of agencies.

Public Elementary/Secondary School. An institution that is linked with an education agency, serves students, and has an administrator. It is possible for

more than one CCD-defined school to exist at a single location (e.g., an elementary and secondary school sharing a building, each with its own principal). One school may also be spread across several locations (e.g., a multiple “storefront” learning center managed by a single administrator).

The CCD classifies schools by type. *Regular schools* provide instruction leading ultimately toward a standard high school diploma; they may also offer a range of specialized services. *Special education* and *vocational schools* have the provision of specialized services as their primary purpose. *Other alternative schools* focus on an instructional area not covered by the first three types (e.g., developing basic language and numeracy skills of adolescents at risk of dropping out of school).

Some schools do not report any students in *membership* (i.e., enrolled on the official CCD reporting day of October 1). This occurs when students are enrolled in more than one school but are reported for only one. For example, students whose instruction is divided between a regular and a vocational school may be reported only in membership for the regular school. In other cases, a school may send the students for which it is responsible to another school for their education—a situation most likely in a small community that does not have sufficient students to warrant keeping a school open every year.

School Locale. Beginning with the 2006-07 CCD files, the locale code methodology was changed from a 1-digit code based on metropolitan statistical areas to a 2-digit code based on urban clusters. American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, the U.S. Virgin Islands, and the Department of Defense Dependents Schools (overseas) were not assigned a locale code because the geographic and governmental structures of these entities do not fit the definitional scheme used to derive the code. There are eight metro-centric locale codes.

The new “urban-centric” locale codes are assigned through a methodology developed by the U.S. Census Bureau’s Population Division in 2005. The urban-centric locale codes apply current geographic concepts to the NCES locale codes used from 1986 through the present. The new urban-centric methodology supplements, and will eventually replace, the older locale code methodology. American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, the U.S. Virgin Islands, and the Department of Defense Dependents Schools (overseas) were not assigned a locale code because the geographic and governmental structures of these entities do not fit the definitional scheme used to derive the code. The Department of Defense Dependents Schools

(domestic) were not assigned locale codes because it is not legal to do so. The new system has 12 urban-centric locale codes.

4. SURVEY DESIGN

Target Population

All public elementary and secondary schools, LEAs, and SEAs throughout the United States, including the District of Columbia, the overseas Department of Defense Dependents Schools, BIE schools, Puerto Rico, and the four outlying areas.

Sample Design

The CCD collects information from the universe of state-level education agencies, except for the Teacher Compensation Survey. The Teacher Compensation Survey is a new survey, and states are participating in it when they are able to report the requested data.

Data Collection and Processing

Through the 2005–06 collection, CCD data were voluntarily obtained from administrative records collected and edited by SEAs during their regular state reporting cycle. In 2006–07, CCD nonfiscal data reporting became mandatory for SEAs. In 2007–08, reporting CCD nonfiscal data to EDFacts, a new data collection system, became mandatory for SEAs.

Reference dates. Most data for the nonfiscal surveys are collected for a particular school year (September through August). The official reference date is October 1st or the closest school day to October 1st. Special education, free-lunch eligibility, and racial/ethnic counts may be taken on December 1st or the closest school day to that date. Student and teacher data are reported for the current school year, whereas through 2005–06, data for high school graduates, other completers, and dropouts reflected the previous year. Fiscal data are for the previous fiscal year; thus, FY 98 data represent the 1997–98 school year.

Data collection. The ways in which CCD data are collected have evolved with the advancement of technology. In the early days of the collection, survey instruments were usually distributed to the states in January. Starting in the 2001–02 collection, downloadable PC software was used. In 2004–05, a web-based data collection application was developed and put into use. A state CCD coordinator, appointed by the Chief State School Officer, is responsible for overseeing the completion of the surveys (often, different coordinators are responsible for the fiscal and the nonfiscal surveys). To assure comparable data across states, NCES provides the CCD coordinator with a set of standard critical definitions for all survey items. In addition, data conferences and training

sessions are held at least yearly. The state's data plan identifies any definitional differences between the state's recordkeeping and the CCD's collection as well as any adjustments made by the state to achieve comparability. Counts across CCD surveys may not be identical, but differences should be consistent and the state is asked to describe the reason for any discrepancy.

NCES provides the state with general information collected during the previous survey on each district and school (e.g., name, address, phone number, locale code, and type of school/district). This information must be verified as correct by the CCD coordinator or recoded with the correct information. The coordinator must also assign appropriate identification codes to new schools and agencies and update the operational status codes for schools and agencies that have closed.

Beginning with the 2005-06 school year, the CCD nonfiscal data have been collected through the U.S. Department of Education's Education Data Exchange Network (EDEN). States report data to EDEN through multiple file groups that fall into various reporting schedule throughout the year. Although states may report data outside the collection period and may revise their reported data at any time in EDEN, NCES extracts the data files from EDEN on the cutoff dates of data submission. The data resubmitted by states after the files were extracted may or may not be included in the CCD final release file.

Data for the CCD fiscal surveys and the TCS are collected by the U.S. Census Bureau. The data are compiled into prescribed formats and submitted by the SEAs. The closing date for the current year's data is the Tuesday following Labor Day. Corrections to submitted fiscal data are accepted until October 1st, however, only corrections that lower a state's current expenditure per pupil are accepted after the "Labor Tuesday" deadline for use in the formula for allocating Title I and other Department of Education funding to state and local school systems.

Editing. Completed surveys undergo comprehensive editing by NCES and the states. Where data are determined to be inconsistent, missing, or out of range, NCES contacts the SEAs for verification. States are given the edit software or are provided with access to the designated website that NCES uses to review data. They are also asked to confirm prepared summaries of the collected information. At this time, the states may revise data collected in the previous survey cycle. NCES examines the data from the 120 largest school districts on a record-by-record basis, setting up fail-safe edit checks to catch unexplained anomalies. In addition, records are processed through a post-edit check to replace blanks and nonmeaningful

zeroes with meaningful responses. After editing, final adjustments for missing data are performed.

Estimation Methods

NCES estimates missing values to improve data comparability across states. Only state-level data are estimated on a regular basis. Missing values in the Public School Universe and Local Agency Universe Surveys are generally left as missing, with a few exceptions. No imputations or adjustments are conducted for state-level data on high school graduates, other high school completer categories, or race/ethnicity.

There are two basic estimation methods: imputation and adjustment. *Imputation* is performed when the missing value for a data item is not reported at all indicating that subtotals and totals containing the category are underreported. Imputation assigns a value to the missing item, and the subtotals and totals containing this item are increased by the amount of the imputation. *Adjustment* corrects a situation in which a value reported for one item contains a value for one or more additional items not reported elsewhere. The original value is reduced by an appropriate amount, which is distributed to the items missing a value. All totals and subtotals are then recalculated. If it is not possible to impute or adjust for a missing value, the item is set to -1 and is counted as "missing."

Every cell in the data file has a companion cell with a flag indicating whether the data contents were reported by the state (R) or placed there by NCES using one of several methodologies: adjustment (A); imputation based on the prior year's data (P); imputation based on a method other than the prior year's data (I); totaling based on the sum of internal or external detail (T); or combining with data provided elsewhere by the state (C).

Estimating state-level nonfiscal data. NCES imputes and adjusts some reported values for student and staff counts at the state level (including the District of Columbia). Imputations for prekindergarten students are performed first, followed by staff imputations and then other adjustments. No imputations or adjustments are made to racial/ethnic data.

Estimating state-level fiscal data. NCES also imputes and adjusts revenue and expenditure data. The federal standard (see National Forum on Education Statistics 2003) is used in the adjustments to distribute expenditure and revenue data. Adjustments are also used to distribute direct state support expenditures to specific objects and functions. In some cases, local revenues from student activities and food services are imputed.

Future Plans

Because it is an ongoing annual survey, the CCD engages in continuous planning with its data users and providers.

5. DATA QUALITY AND COMPARABILITY

The data in the CCD are obtained from the universe of SEAs, which are provided with a common set of definitions for all data items requested. In addition, for the CCD fiscal surveys, NCES provides crosswalk software that converts a state's existing accounting reports to the federal standard, as indicated in *Financial Accounting for Local and State School Systems, 2003 Edition* (National Forum on Education Statistics 2003). This ensures the most comparable and comprehensive information possible across states. As with any survey, however, there are possible sources of error, as described below.

Sampling Error

Because the CCD is a universe survey, its data are not subject to sampling errors.

Nonsampling Error

Coverage error. An NCES report by Owens and Bose (1997), found that overall coverage in the 1994–95 Local Education Agency Universe Survey was 96.2 percent of that in state education directories. “Regular” agencies—those traditionally responsible for providing public education—had almost total coverage. Most coverage discrepancies were attributed to nontraditional agencies that provide special education, vocational education, and other services.

Nonresponse error

Unit nonresponse. The unit of response in the CCD is the SEA. Under current NCES standards, the regular components of the CCD are likely to receive at least partial information from every state, resulting in a 100 percent unit response rate.

Item nonresponse. Any data item missing for one school district is generally missing for other districts in the same state. The following items have higher than normal nonresponse: free-lunch-eligible students by school; nontraditional agencies; and dropouts. Some states assign all ungraded students to one grade and therefore do not report any ungraded students.

Several items have shown marked improvement in response during recent years. Student enrollment was only reported for 80 percent of the districts in 1986–87, but is now available for almost 100 percent. Reports of student race/ethnicity at the school level has increased

from 63 percent in the 1987–88 school year (when first requested) to nearly 100 percent today.

Measurement error. Measurement error typically results from varying interpretations of NCES definitions, differing record keeping systems in the states, and failures to distinguish between zero, missing, and inapplicable in the reporting of data. NCES attempts to minimize these errors by working closely with the state CCD coordinators.

Definitional differences. Although states follow a common set of definitions in their CCD reports, the differences in how states organize education lead to some limitations in the reporting of data, particularly regarding dropouts. CCD definitions appear to be less problematic in the NPEFS, although data on average daily attendance in this survey are not comparable across states. States provide figures for average daily attendance in accordance with state law; NCES provides a definition for states to use in the absence of state law. Because of this lack of comparability, student membership counts from the State Nonfiscal Survey are used as the official state counts.

Because not all states follow the CCD dropout definition and reporting specifications, dropout counts cannot be compared accurately across states. For states that do not comply with the CCD definition, the dropout count is blanked out in the database and considered missing. Currently, there is considerable variation across local, state, and federal data collections on how to define dropouts. The CCD's definition differs from that in other data sources, including the High School and Beyond Study, the National Education Longitudinal Study of 1988, and the Current Population Survey (CPS), conducted by the Census Bureau. Although the collection of dropout information in the CCD is designed to be consistent with procedures in the CPS, differences remain. CCD dropout data are obtained from state administrative records (whereas the CPS obtains this information from a household survey). The CCD includes dropouts in grades 7 through 12 (whereas the CPS includes only grades 10 through 12).

States also vary in the kinds of high school completion credentials on which they collect data. Some states issue a single diploma regardless of the student's course of study. Others award a range of different credentials depending upon whether the student completed the regular curriculum or addressed an individualized set of education goals. Unreported information is shown as missing in CCD data files and published tables unless it is possible to impute or adjust a value (see “Estimation Methods” in section 4 above).

Changes in state reporting practices. The basic characteristics of a school or district do not change frequently. However, a minor change in local or statewide reporting practices (such as two or three coordinators instructing schools to review all of their general information) can have a large impact on the reliability and validity of CCD items. In the 1990–91 school year, a significant proportion (7 percent) of schools, primarily in three states, reported a change in locale code from the prior survey. While this undoubtedly provided better information on school locales in these states, data became less comparable across years. Such changes are rare, however, and tend to be clustered by state and year.

Data Comparability

Most CCD items can be used to assess changes over time by state, district, and school. However, checks of the prevalence and patterns of nonresponse should be performed to assess the feasibility of any analysis. There may also be discontinuities in the data resulting from the introduction of new survey items, changes in state reporting practices, etc., and there may be inconsistencies across reporting levels in the numbers for the same data element (e.g., number of students).

Content changes. As new items are added to the CCD, NCES encourages states to incorporate into their own survey systems the items they do not already collect so that these data will be available in future rounds of the CCD. Over time, this has resulted in fewer missing data cells in each state’s response, thus reducing the need to impute data. Users should keep in mind, however, that while the restructuring of data collection systems can produce more complete and valid data, it can also make data less comparable over time. For example, prior to FY 89, public revenues were aggregated into four categories and expenditures into three functions. Because these broad categories did not provide policymakers with sufficient detail to understand changes in the fiscal conditions of states, the survey was expanded in 1990 to collect detailed data on all public revenues and expenditures within states for regular education in prekindergarten through grade 12.

Comparisons within the CCD. A major goal of the CCD is to provide comparable information across all surveys. The surveys are designed so that the schools in the Public School Universe survey are reflected in the Local Agency Universe survey and so that the data from these surveys are reflected in the State Nonfiscal survey. While counts may not always be equal across reporting levels or even within the same level, differences should be consistent and explainable. For example, counts of students by race/ethnicity in the Public School Universe survey may not always be comparable to student counts by grade because these counts may be taken at different times.

For the most part, the total number of students in a regular district is close to the aggregated number of students in all of the district’s schools. Since 1990, there has typically been agreement between these counts in at least 85 percent of the districts. Membership numbers in the Public School Universe and Local Agency Universe surveys may legitimately differ if (1) there are students served by the district but not accounted to any school (e.g., hospitalized or homebound students); or (2) there are schools operated by the state Board of Education rather than by a local agency. To avoid confusion, NCES publishes the numbers of students and staff from the State Nonfiscal Survey as the official counts for each state.

Teacher counts may also vary across reporting levels. For example, FTE teacher counts are rounded to the nearest hundredth in the Public School Universe survey, but to the nearest whole number in the State Nonfiscal Survey.

6. CONTACT INFORMATION

For content information on the CCD, contact the following individuals:

Program Director:

Marie Stetser
 Phone: 202-502-7356
 E-mail: marie.stetser@ed.gov

Nonfiscal Surveys:

Robert Stillwell
 Phone: (202) 219-7044
 E-mail: robert.stillwell@ed.gov

Nonfiscal data collection and related publications:

Patrick Keaton
 Phone: (202) 502-7386
[E-mail: patrick.keaton@ed.gov](mailto:patrick.keaton@ed.gov)

Fiscal data collection and related publications:

Frank Johnson
 Phone: (202) 502-7362
 E-mail: frank.johnson@ed.gov

Teacher Compensation Survey:

Stephen Cornman
 Phone: (202) 502-7338
 E-mail: stephen.cornman@ed.gov

Mailing Address for All Contacts:

National Center for Education Statistics
 Institute of Education Sciences
 U.S. Department of Education

1990 K Street NW
Washington, DC 20006-5651

7. METHODOLOGY AND EVALUATION REPORTS

General

Sable, J., and Plotts, C. (2009). *Documentation to the NCES Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 2007–08* (NCES 2010-302). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Sable, J., Plotts, C., and Tang, J. (2009). *Documentation to the NCES Common Core of Data Local Education Agency Universe Survey: School Year 2007–08* (NCES 2010-303). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Hamann, T. (2000). *Coverage Evaluation of the 1994–95 Common Core of Data: Public Elementary/Secondary School Universe Survey* (NCES Working Paper 2000-12). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Hoffman, L. (1997). *Customer Service Survey: Common Core of Data Coordinators* (NCES Working Paper 97-15). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Owens, S. and Bose, J. (1997). *Coverage Evaluation of the 1994–95 Common Core of Data: Public Elementary/Secondary Education Agency Universe Survey* (NCES 97-505). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Parrish, T.B., Matsumoto, C.S., and Fowler, W.J. (1995). *Disparities in Public School District Spending 1989–90: A Multivariate, Student-Weighted Analysis, Adjusted for Differences in Geographic Cost of Living and Student Need* (NCES 95-300). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Survey Design

Hamann, T.A. (1999). *Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle* (NCES Working Paper 1999-03). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Chapter 3: Private School Universe Survey (PSS)

1. OVERVIEW

In recognition of the importance of private education, NCES has made the collection of data on private elementary and secondary schools a priority. In 1988, NCES introduced a proposal to develop a private school data collection system that would improve on the irregular collection of private school information dating back to 1890. Since 1989, the U.S. Census Bureau has conducted the biennial Private School Universe Survey (PSS) for NCES. The PSS collects information comparable to that collected on public schools in the Common Core of Data (CCD) (see chapter 2). PSS data are complemented by the more in-depth information collected in the private school sample surveys that are part of the Schools and Staffing Survey (SASS) (see chapter 4). The next PSS data collection will take place during the 2011-12 school year. The next SASS is also planned for the 2011–12 school year.

Purpose

To (1) build an accurate and complete universe of private schools to serve as a sampling frame for NCES surveys of private schools; and (2) generate biennial data on the total number of private schools, teachers, and students.

Components

The PSS consists of a single survey that is completed by administrative personnel in private schools. An early estimates survey designed to allow early reporting of key statistics was discontinued after the 1992–93 school year.

Private School Universe Survey. This survey collects data on private elementary and secondary schools, including religious orientation, level of school, length of school year, length of school day, total enrollment (K–12), race/ethnicity of students, number of high school graduates, number of teachers employed, program emphasis, and existence and type of kindergarten program.

Periodicity

Biennial. The next PSS will be administered in 2011-12 and every 2 years thereafter. Earlier surveys were conducted in 1989–90, 1991–92, 1993–94, 1995–96, 1997–98, 1999–2000, 2001–02, 2003–04, 2005–06, and 2009-10.

2. USES OF DATA

The PSS produces private school data similar to that produced for public schools in the CCD. Profiles of private education providers can be developed from PSS data to address a variety of policy- and research-relevant issues, including the growth of religiously affiliated schools, the number of private high school graduates, the

BIENNIAL SURVEY OF THE UNIVERSE OF PRIVATE SCHOOLS

PSS collects data on:

- Student enrollment
- Teaching staff
- High school graduates
- School religious affiliation

length of the school year for various private schools, and the number of private school students and teachers.

3. KEY CONCEPTS

Some key concepts related to the PSS are described below.

Private School. A school that is not supported primarily by public funds. It must provide classroom instruction for one or more of grades K–12 (or comparable ungraded levels) and have one or more teachers. Organizations or institutions that provide support for home schooling but do not offer classroom instruction for students are not included. Private schools are assigned to one of three major categories and, within each major category, to one of three subcategories:

- *Catholic:* parochial, diocesan, private;
- *Other religious:* affiliated with a conservative Christian school association, affiliated with a national denomination, unaffiliated; and
- *Nonsectarian:* regular program emphasis, special program emphasis, special education.

Schools with kindergarten, but no grade higher than kindergarten, are referred to as *kindergarten-terminal (K-terminal)* schools; these schools were first included in the 1995–96 PSS. Schools meeting the pre-1995 definition of a private school (i.e., including any of grades 1–12) are referred to as *traditional* schools.

Elementary School. A school with one or more of grades K–6 and no grade higher than grade 8. For example, schools with grades K–6, 1–3, or 6–8 are classified as elementary schools.

Secondary School. A school with one or more of grades 7–12 and no grade lower than grade 7. For example, schools with grades 9–12, 7–8, 10–12, or 7–9 are classified as secondary schools.

Combined School. A school with one or more of grades K–6 and one or more of grades 9–12. For example, schools with grades K–12, 6–12, 6–9, or 1–12 are classified as combined schools. Schools in which all students are ungraded (i.e., not classified by standard grade levels) are also classified as combined.

Teacher. Any full- or part-time teacher whose school reports that his or her assignment is teaching in any of grades K–12.

4. SURVEY DESIGN

Target Population

All private schools in the United States that meet the NCES definition. The PSS universe consists of a diverse population of schools. It includes both schools with a religious orientation (e.g., Catholic, Lutheran, or Jewish) and nonsectarian schools with programs ranging from regular to special emphasis and special education.

Sample Design

NCES uses a dual-frame approach for building its private school universe. The primary source of the PSS universe is a *list frame* containing most private schools in the country. The list frame is supplemented by an *area frame*, which contains additional schools identified during a search of randomly selected geographic areas around the country. The two frames are used together to estimate the population of private schools in the United States. Since documentation for the 2009–10 PSS has not been completed, these descriptions are for the 2007–08 PSS.

List frame. In an effort to ensure a complete population list of all private elementary and secondary schools in the United States, NCES updates the list frame every 2 years in preparation for the next PSS administration. The list frame was initially developed for the 1989–90 survey. The list is updated periodically by matching it with lists provided by nationwide private school associations, state departments of education, and other national private school guides and sources.

The basis of the current survey's list frame is the previous PSS. In order to expand coverage to include private schools founded since the previous survey, NCES requests lists of schools from the 50 states and the District of Columbia in advance of each survey administration. Requests are made to state education departments, as well as to other departments, such as health or recreation. NCES also collects membership lists from about 29 private school associations and religious denominations. Schools on the state and association lists are compared to the base list, and any school not matching a school on the base list is added to the universe list.

Prior to the 1995–96 survey, only schools that included at least one of grades 1–12 were included in the PSS (now referred to as *traditional schools*). As of 1995–96, the PSS has also collected data from K-terminal schools. NCES also removed from the PSS eligibility criteria the requirements that a school have 160 days in the school year and 4 hours per day during which classes are conducted.

In 2007, a separate list-building operation (Early Childhood Operation) was conducted to identify K-terminal schools. Requests for lists of programs that might include a kindergarten were made to sources, other than state departments of education, in all 50 states and the District of Columbia, including state departments of health or recreation; state child care licensing agencies; and child care referral agencies. In 2007, some 24 of these early childhood lists were received, and 19 were processed (due to resource constraints, not all of the lists were processed).

Schools on private school association membership lists, the state lists, and the early childhood lists were compared to the base list, and any school that did not match a school on the base list was added to the universe list. Additionally, questionnaires were sent out to programs identified in the 2005–06 PSS as prekindergarten only. This procedure was done in case any of these programs included at least a kindergarten in the 2007–08 school year. A total of 37,275 schools (unweighted) were included in the 2007–08 list frame.

Area frame. The list frame is supplemented by an area frame, which contains additional private schools identified during a search of telephone books and other sources in randomly selected geographic areas around the country. The area frame search is conducted by the Bureau of the Census. Each area's list is created from a set of predetermined sources within that area and then matched against the updated list frame universe to identify schools missing from the updated list frame.

The United States is divided into 2,062 primary sampling units (PSUs), each consisting of a single county, independent city, or cluster of geographically contiguous areas. The eight PSUs with the highest private school enrollment in the 2000 census populations greater than 1.7 million were selected with certainty for the private school survey. In addition to these certainty PSUs, the area frame consists of two sets of sample PSUs: (1) a 50 percent subsample (overlap) of the area frame sample PSUs from the previous PSS, to maintain a reasonable level of reliability in estimates of change, and (2) a sample of PSUs selected independently from the previous PSS sample (nonoverlap PSUs). A minimum of two nonoverlap PSUs are allocated to each of the 16 strata, which are defined by (1) four Census regions (Northeast, Midwest, South, or West); (2) metro/nonmetro status (two levels); and (3) whether the PSU's percentage of private school enrollment exceeds the median percentage of private enrollment of the other PSUs in the census region/metro status strata (two levels). Within a stratum, the sample PSUs are selected with probability proportional to the square root of the population in each of the PSUs.

A total of 124 distinct PSUs (162 counties) were in the 2007–08 PSS area frame sample. Within each of these PSUs, the Census Bureau attempted to find all eligible private schools. A block-by-block listing of all private schools in a sample of PSUs was not attempted. Rather, regional office field staff created the frame by using such sources as the yellow pages, local Catholic dioceses, religious institutions, local education agencies, and local government offices. Once the area search lists were constructed, they were matched with the NCES private school universe list. Schools that did not match the universe list were deleted from the area frame. A total of 1,872 schools (unweighted) were added to the universe from the area frame.

Due to differences in methodology and definition, the results of the 1993–94 and subsequent area search frames are not strictly comparable to the results of earlier years. Prior to 1993, an initial eligibility screening was performed by telephone for area frame schools before the questionnaire was mailed out. Ineligible schools were declared out of scope at that time, and eligible schools were either interviewed by telephone or sent a questionnaire. In the 1993–94 PSS, screener questions were added to the survey instrument to determine eligibility. Ineligible schools were not eliminated until the questionnaires were returned. In the 1995–96 PSS, all area frame schools were placed in the telephone follow-up phase of the PSS, and ineligible schools were again eliminated based on responses to screener questions.

Data Collection and Processing

The data collection phase consists of (1) a mailout/mailback stage; and (2) a telephone follow-up stage. The U.S. Census Bureau is the collection agent.

Reference dates. The official reference date for reporting PSS information is October 1.

Data collection. In October of the survey year, the Census Bureau mails PSS questionnaires to the private schools. (Data collection for the 2007–08 PSS coincided with the data collection phase of the private school component of the 2007–08 SASS: the private schools selected for SASS were excluded from the PSS, and the schools selected for SASS received a SASS private school questionnaire only, while the remaining private schools were sent a PSS questionnaire. The PSS questionnaire used the same wording as the SASS questionnaire, but contained only a subset of the SASS questionnaire items. After data collection, the data for the SASS cases were merged into the PSS universe.) If no response is received within a month, a second questionnaire is mailed. Reminder postcards are sent 1 week after each questionnaire mailout. Three to 4 months after the initial mailout, the Census Bureau begins telephone follow-up of schools that have not responded to either

mailout; the schools from the area frame operation are added at this time. Interviewing takes place at the Census Bureau's computer-assisted telephone interviewing (CATI) facilities. For schools that cannot be contacted by telephone, additional follow-up is conducted in the Census Bureau's regional offices.

Editing. Most of the mailback questionnaires are scanned; those that must be keyed are 100 percent key-verified. For data collected during the telephone follow-up phase, preliminary quality assurance and editing checks take place at the time of the interview. The data collection instrument is designed to alert interviewers to inconsistencies reported by the respondent so that any necessary corrections can be made at this time. Data from the CATI facilities are transmitted to Census headquarters for further processing where they undergo extensive editing, including:

- range checks to eliminate out-of-range entries;
- consistency edits to compare data in different fields for consistency;
- blanking edits to verify that skip patterns on the questionnaire were followed; and
- interview status recodes (ISRs), performed prior to the weighting process, to assign the final interview status to the records (i.e., interview, noninterview, or out-of-scope).

Estimation Methods

Weighting adjusts the number of schools in the area frame sample up to a fully representative number of schools missing from the list frame and adjusts the survey data from both the area and list components for school nonresponse. Imputation is used to compensate for item nonresponse.

Weighting. PSS data from the area frame component are weighted to reflect the sampling rates (probability of selection) in the PSUs. Survey data from both the list and area frame components are adjusted for school nonresponse. This represents a departure from procedures used in the 1989–90 survey, which adjusted for total nonresponse (i.e., school nonresponse) and for partial nonresponse associated with four specific PSS data elements. Since 1991, only one weight has been required, due to a newly developed and complex imputation process used to compensate for item nonresponse. When estimates are produced for schools and other data elements, the same PSS school weight should be used. A brief description of the components comprising the PSS weight follows:

W_i , the PSS weight for all data items for the i^{th} school, is

$$W_i = BW_i \times NR_c$$

where BW_i is the **base weight**, or the inverse of the selection probability for school i ($BW_i = 1$ for list frame schools; $BW_i =$ the inverse of the PSU probability of selection for area frame schools), and

NR_c is the **nonresponse adjustment factor**, or weighted ratio of the sum of the in-scope schools to the sum of the in-scope responding schools in cell c , using BW_i as the weight.

The cells used to compute the nonresponse adjustment are defined differently for list-frame and area-frame schools. In 2007–08 PSS, for schools in the list frame, the cells were defined by affiliation, urbanicity type, grade level, region, and enrollment. The nonresponse adjustment cells for area frame schools were defined by certainty/noncertainty PSU status, three-level typology (Catholic, Other religious, Nonsectarian), and grade level.

If the number of schools in a cell was less than 15 or the nonresponse adjustment factor was greater than 1.5, then that cell was collapsed into a similar cell. The cells for traditional schools from the list frame were collapsed within enrollment category, urbanicity type, grade level, and census region. Cells for K-terminal schools from the list frame were collapsed within enrollment category, urbanicity type, region (if applicable), and affiliation. Cells for traditional schools from the area frame were collapsed within grade level and then within three-level typology. Cells for K-terminal schools from the area frame were collapsed within three-level typology.

Imputation. Since the 1991–92 PSS, imputation has been used to compensate for item nonresponse in records classified as interviews (i.e., required items are completed). All items that are missing data are imputed. The first survey, the 1989–90 PSS, used weighting adjustments for both interviews and noninterviews.

Imputation occurs in two stages. The first-stage (internal) process uses data from other items for the same school in the current PSS and data from the previous PSS. If an item cannot be imputed during the first-stage process, it is imputed during the second stage. The second-stage (donor) process uses a hot-deck imputation methodology that extracts data from the record for a reporting school (donor) similar to the nonrespondent school. All records (donors and nonrespondents) in the file are sorted by variables that describe certain characteristics of the schools, such as

school type, affiliation, school level, enrollment, and urbanicity.

For a few items, entries are clerically imputed. The data record, sample file record, and the questionnaire are reviewed, and an entry consistent with the information from those sources is imputed. This procedure is used when: (1) no suitable donor is found, (2) the computer method produces an imputed entry that is unacceptable, and (3) the nature of the item requires an actual review of the data rather than a computer-generated value.

Recent Changes

Several changes to the questionnaire have been introduced in the previous PSS cycles. In the 1993–94 PSS, three major revisions were made. First, a new design was implemented to facilitate respondent reporting by clearly indicating skip patterns through the use of arrows as well as words and by minimizing the number of questions asked on each page. Second, content on prekindergarten programs was expanded to collect the type of prekindergarten program in addition to the prekindergarten student and teacher counts requested in earlier surveys (these data were collected as a part of a separate Census Bureau initiative and are not included in PSS reports). Third, data on the racial/ethnic makeup of the school's student body were collected for the first time.

Modifications made to the 1995–96 PSS included adding nursery and prekindergarten, transitional kindergarten, and transitional first-grade enrollment counts to the enrollment item. Questions on the length of the school day and number of days per week for kindergarten, transitional kindergarten, and transitional first grade were also added. "Early childhood program/day care center" was added as a category for type of school. The 1993–94 PSS questionnaire items concerning types of prekindergarten programs and the number of prekindergarten teachers were deleted.

In the 1997–98 PSS, the following items were added to the survey instrument: (1) whether or not the school is coeducational (if yes, the number of male students; if no, whether the school is all female or all male); and (2) whether or not the school has a library or library media center.

There were few changes in the 1999–2000 PSS. One religious affiliation—Church of God in Christ—was added, and three associations were added—Association of Christian Teachers and Schools, National Coalition of Girls' Schools, and state or regional independent school associations. The item that previously collected data on the number of graduates that applied to 2-year or 4-year colleges was changed to collect data on the percentage of graduates who went on to attend three types of schools: 2-year colleges, 4-year colleges, and

technical or other specialized schools. There also was a minor change in the definition of community type. Beginning with the 1999–2000 PSS, schools that were "rural within a Metropolitan Statistical Area" were included in the "Rural/small town" community type, while prior to the 1999–2000 PSS they were included in the "Urban fringe/large town" community type.

The 2001–02 PSS questionnaire content was relatively unchanged from the 1999–2000. One question was added to item 2 (the screener item)—"Is the school named on the front of this questionnaire located in the United States?" This question was added to facilitate the exclusion of schools from the PSS that were located outside of the United States, but had been added during the list building or area search because the school had an office with an address in the United States.

Additionally, in order to test the feasibility and benefits of collecting PSS data over the Internet, the 2001–02 PSS included an Internet response option test. The final response rate for Internet submissions was 15.4 percent for schools that received the option (5.1 percent of all schools).

Changes made to the 2003–04 PSS were minor and involved frame creation methodology, data collection procedures, and weighting procedures. For example, whereas in the 2001–02 PSS, the base weight for area frame schools was equal to the inverse of the probability of selecting the PSU in which the school resided, in the 2003–04 PSS, the base weight for area frame schools also contained a nonunitary subsampling factor for schools named solely in non-Roman Catholic religious institution lists.

Caution, however, should be used in comparing 2003–04 PSS community type estimates to those of previous years. Although the definition of community type remained unchanged, the 2003–04 PSS community types are based on the Consolidated Statistical Area/Core-Based Statistical Area rather than on the Standard Metropolitan Statistical Area/Metropolitan Statistical Area, which was used prior to the 2003–04 PSS. Also, community type is based on 2000 census data; prior to the 2003–04 PSS, community type was based on 1990 census information.

There were few changes in the 2005–06 PSS. One religious affiliation—Church of the Nazarene—was added. Also, the 2005–06 PSS used the new 12-level urban-centric locale codes, rather than the 8-level locale codes based on the Core-Based Statistical Area.

There was one change in the 2007–08 PSS. In 2005–06, non-Roman Catholic religious institutions were contacted during the area-frame operation while in 2007–08 they were not.

Future Plans

The PSS will continue as a biennial survey.

5. DATA QUALITY AND COMPARABILITY

Sampling Error

Only the area frame contributes to the standard error in the PSS. The list frame component of the standard error is always 0. Estimates of standard errors are computed using half-sample replication.

Because the area frame sample of PSUs is small (125 out of a total of approximately 2,000 eligible PSUs), there is a potential for unstable estimates of standard errors. This is particularly true when the domain of interest is small and there may not be enough information to compute a standard error. Stabilizing the standard error estimate given the level of detail of the PSS estimates would require a much larger PSU sample. The current area frame is designed to produce regional estimates.

Nonsampling Error

Coverage error. Undercoverage in the list and area frames is one possible source of nonsampling error. Because the PSS uses a dual-frame approach, it is possible to estimate the coverage, or completeness, of the PSS. A capture-recapture methodology is used to estimate the number of private schools in the United States and to estimate the coverage of private schools. In the 2003–04 PSS, the conservative coverage rate for traditional private schools was equal to 96 percent; for K-terminal private schools, it was equal to 85 percent. In the 2005–06 PSS, the overall coverage rate was 98 percent. In the 2007–08 PSS, the conservative coverage rate for traditional private schools was equal to 96 percent; for K-terminal private schools, it was equal to 93 percent.

A study comparing the quality of PSS frame coverage to that of the commercial Quality Education Data database of schools is discussed in Lee, Burke, and Rust (2000).

Nonresponse error. There are two types of nonresponse error: unit nonresponse and item nonresponse.

Unit nonresponse. In the 2007–08 PSS, the survey data from the area frame component were weighted to reflect the sampling rates (probability of selection) of the PSUs. Survey data from both the list and area frame components were adjusted for school nonresponse. There were 28,450 interviews and 2,527 cases that were noninterviews. After weighting the area frame component, these became 30,748 interviews and 2,992

noninterviews—the weighted response rate was 91 percent. In the 2005–06 PSS, the survey data from the area frame component were weighted to reflect the sampling rates (probability of selection) of the PSUs. Survey data from both the list and area frame components were adjusted for school nonresponse. There were 29,784 interviews and 1,867 cases that were noninterviews. After weighting the area frame component, these became 32,865 interviews and 2,159 noninterviews—the weighted response rate was 94 percent. In the 2003–04 PSS, of the 41,184 schools included (both traditional and K-terminal), some 9,336 cases were considered out-of-scope (that is, not eligible for the PSS). A total of 30,071 private schools completed a PSS interview, while 1,777 schools refused to participate, resulting in an overall unweighted response rate of 94 percent. When the area frame schools were weighted by the inverse of the probability of selection, the weighted response rate was 94 percent as well. In the 2001–02 PSS, the weighted response rate for traditional schools was 95 percent (96 percent unweighted); for K-terminal schools, the response rates were 97 and 96 percent, respectively. In 1999–2000, both the weighted and unweighted response rates were 93 percent for traditional schools; they were 99 and 98 percent, respectively, for K-terminal schools.

Item nonresponse. In the 2007–08 PSS, all of the weighted response rates were greater than 85 percent. The weighted item response rates for all but one variable—the percentage of graduates who went to 2-year colleges—were greater than 85 percent in 2005–06. In the 2003–04 PSS, all of the weighted response rates were greater than 85 percent. In the 2001–02 PSS, for traditional schools, all but three items had weighted response rates greater than 90 percent. The three lower rates (ranging from 77.5 percent to 86.3 percent) pertained to the percentage of graduates who went to 4-year colleges, 2-year colleges, and technical or other specialized schools. Values for items with missing data were imputed to compensate for item nonresponse.

Measurement error. NCES seeks to minimize measurement error by developing survey content in consultation with representatives of private school associations, reviewing extensively the questionnaire and instructions before distribution, requiring that the data that are not scanned are 100 percent key-verified, and processing the survey data through a comprehensive series of edits to verify accuracy and consistency.

Intersurvey Consistency in NCES Private School Surveys

The PSS and the private school component of SASS were fielded in the same school year for the first time in 1993–94. Even though these two surveys measure some of the same variables (schools, teachers, and

students), the 1993–94 results were not in agreement due to sampling and other errors. PSS results are likely to be the more accurate since the PSS serves as the sampling frame for the SASS private school component (a sample of around 3,000 schools). Special methodological studies of these two surveys have been done, including comparisons among statistical and computational procedures aimed at achieving consistency between the estimates of private schools, private school teachers, and private school students in the 1993–94 PSS and in the 1993–94 SASS—see Scheuren and Li (1995, 1996).

Data Comparability

While changes to survey design and content generally result in improved data quality, they also impact the comparability of data over time. Recent changes to the PSS and to the comparability of PSS data (both within the PSS itself and with other data sources) are discussed below.

Design change. Changes in the survey design of the 1995–96 PSS resulted in an increased number of private schools in the survey population. First, seven new association lists were obtained, adding 512 new schools to the list frame. In previous years, the area frame was relied upon to include these schools. Second, the area search results were not strictly comparable to those in previous years due to procedural differences. The 1995–96 PSS was the first survey to verify the control of schools marked as public in the screener item. Final determination of school control was based on a review of the school’s name and other identifying information. As a result, several schools that had been marked as public (but which were obviously private) were added back into the PSS. They were counted as interviews if the required data were provided or as noninterviews if the required data were missing. Third, the eligibility criteria for the PSS were changed to no longer require schools to have 160 days in the school year or to conduct classes for at least 4 hours per day. Fourth, the PSS definition of a school was expanded to include programs where kindergarten is the highest grade (K-terminal schools). Additional lists of programs that might have a kindergarten were requested from nontraditional sources, and the area search was expanded to search for programs with a kindergarten. Some schools meeting the traditional PSS definition of a school (any of grades 1–12 or comparable ungraded levels) were discovered in these lists. When added to the PSS, these schools also increased the estimates of traditional schools.

Note that even when the population of schools is about the same from one survey to the next, it may represent a different set of schools. For example, the number of schools was around 27,000 in both 1997–98 and 1999–2000, although about 1,700 schools were added to the PSS universe in 1999–2000. This suggests that a nearly

equal number of schools dropped out of the universe between 1997–98 and 1999–2000. Comparisons of the 1999–2000 PSS private school estimates with those from the 2001–02 PSS, however, show an overall increase in the number of private schools between 1999–2000 and 2001–02 (to about 29,000).

Questionnaire changes. Several modifications have been made to the format and content of the PSS questionnaire since 1991–92. A number of items were added (including race/ethnicity of students), and some items were deleted or modified.

Comparisons within the PSS. The estimated number of schools decreased between 2005–06 and 2007–08 (by 1,314 schools). The estimated number of private students and full-time-equivalent (FTE) teachers in 2007–08 were not statistically different from those of 2005–06. The estimated number of private schools and students decreased between the 2001–02 and 2003–04 PSS data collections (by 889 schools and 218,741 students). The estimated number of FTE teachers in 2003–04 was not statistically different from that in 2001–02. Comparisons of the 2001–02 PSS estimates with those from previous PSS data collections show increases in the number of private schools, students, and teachers between 1999–2000 and 2001–02. Comparisons of the 1999–2000 PSS estimates with those from previous surveys show no significant change in the estimated number of private schools; however, they do indicate an increase in the estimated number of private school teachers and students.

Comparisons with the Current Population Survey. A comparison of the PSS estimate of K–12 students enrolled in all private schools (traditional and K-terminal) with the household survey estimate from the 2007 October Supplement of the Current Population Survey (CPS) (U.S. Census Bureau 2006) shows that the PSS estimate of 5,072,451 does statistically differ from the CPS estimate of the number of private school students in grades kindergarten through 12 in October 2007 of 4,817,000. A comparison of the 2003–04 PSS estimate of K–12 students enrolled in all private schools (traditional and K-terminal) with the household survey estimate from the 2003 October Supplement to the CPS shows that the PSS estimate of 5,212,992 students is not statistically different from the CPS estimate of 5,259,000 students (U.S. Census Bureau 2005). A comparison of the 2001–02 PSS estimate of K–12 students enrolled in all private schools (traditional and K-terminal) with the household survey estimate from the October 2001 CPS shows that the PSS estimate of 5,439,925 is higher than the CPS estimate of 5,164,000; the 95 percent confidence interval of the PSS estimate ranges from 5,383,898 to 5,495,952 students, while that of the CPS estimate ranges from 4,956,000 to 5,372,000 students. In the 1995–96 school year, the PSS and CPS estimates did

not differ significantly; in 1997–98, the PSS estimate was higher than the CPS estimate; and, in 1999–2000, the PSS estimate was lower than the CPS estimate. Comparisons between CPS and PSS enrollment estimates for earlier years are not as informative since, prior to 1995–96, the PSS estimates did not include the kindergarten enrollment from K-terminal schools, whereas the CPS has always included it.

Comparisons with National Catholic Educational Association data. Comparisons of the PSS estimates for Catholic schools, students, and FTE teachers (traditional schools) with the National Catholic Educational Association (NCEA) (National Catholic Educational Association 2008) data for the 2007–08 school year show differences in the school (7,507 versus 7,378), student (2,156,173 versus 2,270,913) and FTE teacher counts (146,627 versus 160,075) between PSS and NCEA, respectively. Comparisons of the PSS estimates for Catholic schools, students, and FTE teachers with the NCEA data for the 2003–04 school year show differences in the number of students (2,365,220 vs. 2,484,252) and FTE teachers (152,611 vs. 162,337) between PSS and NCEA, respectively. The difference between the PSS estimate of 7,919 Catholic schools and the NCEA count of 7,955 schools is not statistically significant. The survey methodologies used by NCES and NCEA are quite different; NCES surveys private schools directly, while NCEA surveys archdiocesan and diocesan offices of education and some state Catholic conferences. The NCEA and PSS computations of full-time equivalents differ in the weight assigned to part-time teachers; thus, the PSS and NCEA counts of FTE teachers are not strictly comparable.

For the 2001–02 school year, comparisons of the PSS estimate for Catholic schools with the NCEA data show differences in the school and student counts. The NCEA count of 8,000 schools is below the lower limit of the 95 percent confidence interval of the PSS estimate of Catholic schools (which ranges from 8,112 to 8,302). The NCEA K–12 student count of 2,553,277 is higher than the upper limit of the 95 percent confidence interval of the PSS estimate of Catholic students (which ranges from 2,492,773 to 2,538,274). Both the NCEA teacher count of 163,004 and the PSS estimate of 155,514 include part- and full-time teachers in the computation of full-time equivalents (the 95 percent confidence interval of the PSS estimate ranges from 153,902 to 157,126).

NCES publication criteria for the PSS. NCES criteria for the publication of an estimate are dependent on the type of survey—sample or universe. To publish an estimate for a sample survey, at least 30 cases must be used in developing the estimate. For a universe survey, a minimum of three cases must be used. The PSS includes both types of surveys: (1) a sample survey of

PSUs (area frame) that collects data on schools not in the list frame (the number of PSUs changes for each administration); and (2) a complete census of schools belonging to the list frame. NCES has established a rule that published PSS estimates must be based on at least 15 schools. If the estimate satisfies this criterion and the coefficient of variation (standard error/estimate) is greater than 25 percent, the estimate is identified as having a large coefficient of variation and the reader is referred to a table of standard errors.

6. CONTACT INFORMATION

For content information on the PSS, contact:

Stephen Broughman
Phone: (202) 502-7315
E-mail: stephen.broughman@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Broughman, S.P. (1996). *Private School Universe Survey, 1993–94* (NCES 96-143). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Broughman, S.P., and Colaciello, L.A. (1998). *Private School Universe Survey, 1995–96* (NCES 98-229). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Broughman, S.P., and Colaciello, L.A. (1999). *Private School Universe Survey, 1997–98* (NCES 1999-319). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Broughman, S.P., and Colaciello, L.A. (2001). *Private School Universe Survey, 1999–2000* (NCES 2001-330). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Broughman, S.P., Gerald, E., Bynum, L.T., and Stoner, K. (1994). *Private School Universe Survey, 1991–92* (NCES 94-350). National Center

for Education Statistics, U.S. Department of Education. Washington, DC.

Broughman, S.P., and Pugh, K.W. (2004). *Characteristics of Private Schools in the United States: Results From the 2001–2002 Private School Universe Survey* (NCES 2005-305). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Broughman, S.P., and Swaim, N.L. (2006). *Characteristics of Private Schools in the United States: Results From the 2003–2004 Private School Universe Survey* (NCES 2006-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Broughman, S.P., Swaim, N.L., and Keaton, P.W. (2008). *Characteristics of Private Schools in the United States: Results From the 2005–2006 Private School Universe Survey* (NCES 2008-315). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Broughman, S.P., Swaim, N.L., and Keaton, P.W. (2009). *Characteristics of Private Schools in the United States: Results From the 2007–08 Private School Universe Survey* (NCES 2009-313). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Gerald, E., McMillen, M., and Kaufman, S. (1992). *Private School Universe Survey, 1989–90* (NCES 93-122). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

National Catholic Educational Association. (2008). *United States Catholic Elementary and Secondary Schools, 2007–08: Annual Statistical Report on Schools, Enrollment, and Staffing*. Washington, DC: Author.

U.S. Census Bureau. (2005). School Enrollment-Social and Economic Characteristics of Students: October 2003. Retrieved March 24, 2005, from <http://www.census.gov/population/www/socdemo/school/spw2003.html>.

U.S. Census Bureau. (2006). School Enrollment-Social and Economic Characteristics of Students: October 2005. Retrieved March 24, 2009, from <http://www.census.gov/population/www/socdemo/school/cps2005.html>.

Survey Design

McMillen, M., and Benson, P. (1992). *Diversity of Private Schools* (NCES 92-082). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Scheuren, F., and Li, B. (1995). *Intersurvey Consistency in NCES Private School Surveys* (NCES Working Paper 95-16). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Scheuren, F., and Li, B. (1996). *Intersurvey Consistency in NCES Private School Surveys for 1993–94* (NCES Working Paper 96-27). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Broughman, S., Tourkin, S., Swaim, N., Hryczaniuk, C., Guan, X., Peterson, J., Parmer, R., Abramson, R., Zhang, B., Andriani, S., & Walter, E. (2009). *2007–08 Private School Universe Survey (PSS) Data File User's Manual and Survey Documentation* (NCES 2009-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Jackson, B.J., and Frazier, R.L. (1996). *Improving the Coverage of Private Elementary-Secondary Schools* (NCES Working Paper 96-26). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Jackson, B.J., Johnson, N.R., and Frazier, R.L. (1997). Improving the Coverage of Private Elementary-Secondary Schools. In *Selected Papers on the Schools and Staffing Survey: Papers Presented at the 1997 Meeting of the American Statistical Association* (NCES Working Paper 97-41). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Lee, H., Burke, J., and Rust K. (2000). Evaluating the Coverage of the U.S. National Center for Education Statistics' Public and Private School Frames Using Data From the National Assessment of Educational Progress. In *Proceedings of the Second International Conference on Establishment Surveys: Survey Methods for Businesses, Farms, and Institutions* (pp. 89–98). Arlington, VA: American Statistical Association.

Tourkin, S., Swaim, N., Guan, X., Peterson, J., Schwanz, D., Warner, T., Parmer, R., Wilson, M., Broughman, S., and Walter, E. (2008). *2003–04 Private School Universe Survey (PSS) Data File User's Manual and Survey Documentation* (NCES

2008-314). National Center for Education Statistics,
Institute of Education Sciences, U.S. Department of
Education. Washington, DC.

Chapter 4: Schools and Staffing Survey (SASS)

1. OVERVIEW

The NCES Schools and Staffing Survey (SASS) provides data on public and private schools, principals, school districts, and teachers. SASS gathers information about many topics, including various characteristics of elementary and secondary students, some of the professional and paraprofessional staff who serve them, the programs offered by schools, principals' and teachers' perceptions of school climate and problems in their schools, teacher compensation, and district hiring practices. SASS is a unified set of surveys that facilitates comparisons between public and private schools and allows linkages of teacher, school, school district, and principal data. First conducted in school year 1987–88, SASS has been conducted six times, most recently in school year 2007–08.

Purpose

The purpose of SASS is to collect the information necessary for a complete picture of American elementary and secondary education. SASS is designed to provide national estimates of public elementary, secondary, and combined schools and teachers; state estimates of public elementary and secondary schools and teachers; and estimates for private schools; teachers and principals at the national level; and by private school affiliation. The SASS questionnaires were revised for the 2003–04 and the 2007–08 administrations, with the addition of new items about teachers' career paths, parental involvement, school safety, and institutional support for information literacy. The questionnaires continued to measure the same five policy issues: teacher shortage and demand; characteristics of elementary and secondary teachers; teacher workplace conditions; characteristics of school principals; and school programs and policies.

Core Components

SASS consists of four core components administered to districts, schools, principals, and teachers. The district questionnaire is sent to a sample of public school districts. The school questionnaire is sent to a sample of public schools and private schools, as well as all charter schools in operation as of 1998–99, and all schools operated by the Bureau of Indian Education (BIE) or American Indian/Alaska Native tribes. The principal and teacher questionnaires are sent to a sample of principals and teachers working at the schools that receive the school questionnaire. (The Teacher Follow-up Survey is a fifth component of SASS and is covered in chapter 5.)

School District Survey (formerly the Teacher Demand and Shortage Survey). The questionnaire for this survey is mailed to each sampled local education agency (LEA). The respondents are contact people identified by LEA personnel.

SAMPLE SURVEY OF PUBLIC, PRIVATE, CHARTER, AND BIE SCHOOLS

SASS collects data on:

- School districts
- Principals
- Schools
- Teachers
- Library media centers

If no contact person is identified, the questionnaire is addressed to “Research Director.” The School District Questionnaire consists of items about student enrollment, number of teachers, teacher recruitment and hiring practices, teacher dismissals, existence of a teacher union, length of the contract year, teacher compensation, school choice, magnet programs, graduation requirements, oversight of home-schooled students and charter schools, use of school performance reports, migrant education, and professional development for teachers and administrators. Some items that appeared previously have been dropped, such as those that collected layoff data and counts of students by grade level (the latter are available through the NCES Common Core of Data [CCD]). In the 2003–04 administration, new topics, including principal hiring practices and instructional aide hiring practices, were added to the questionnaire. In the 2007–08 administration, items on district performance, teacher tenure and dismissal, principal salary, length of the contract year for teachers, and type of retirement benefits for teachers were added or revised.

The School District Questionnaire is mailed only to public school districts. Independent public charter schools, BIE-funded schools, and schools that are the only school in the district are given the School Questionnaire (with district items), not the School District Questionnaire. The School Questionnaire (with district items) includes all of the items included in the School Questionnaire as well as selected items from the School District Questionnaire. The applicable items for private schools appear in the Private School Questionnaire.

School Principal Survey (formerly the School Administrator Survey). The questionnaire for this survey collects information about principal/school head demographic characteristics, training, experience, salary, and judgments about the seriousness of school problems. Information is also obtained on professional development opportunities for teachers and principals, teacher performance, barriers to dismissal of underperforming teachers, school climate and safety, parent/guardian participation in school events, and attitudes about educational goals and school governance. The 2007–08 questionnaire appeared in two versions: one for principals or heads of public schools and one for heads of private schools. The two versions contain minor variations in phrasing to reflect differences between public and private schools in governing bodies and position titles in schools. Items on experience prior to becoming a principal, teacher and school performance, and time allocation for

students during the week were added or revised in the 2007–08 questionnaire.

School Survey. The questionnaires for this survey are sent to public schools, private schools, BIE schools, and charter schools. Private schools receive the Private School Questionnaire, while BIE schools and charter schools receive the School Questionnaire (with district items), described separately below. As in 2003–04, the 2007–08 data collection for the private school component of SASS coincided with the administration of the NCES Private School Universe Survey (PSS). Since both PSS and SASS were administered in 2007–08, to reduce respondent burden, the private schools in the SASS sample were not sent a PSS questionnaire. Instead, the PSS items appeared in the SASS Private School Questionnaire. (See chapter 3 for a complete description of PSS.)

The School Questionnaire is addressed to “Principal,” although the respondent can be any knowledgeable school staff member (e.g., vice principal, head teacher, or school secretary). Items cover grades offered, student attendance and enrollment, staffing patterns, teaching vacancies, high school graduation rates, programs and services offered, curriculum, and college application rates. The Private School Questionnaire also includes items from the School District Questionnaire that are applicable to private schools. The 2007–08 collection included items on the beginning time of students’ school day; length of the school year for students; school websites; and math, reading, or science specialist assignments.

School Questionnaire (with district items). The purpose of the questionnaire (which was also referred to as the Unified School Questionnaire in the 2003–04 SASS) was to obtain information about schools, such as grades offered, number of students enrolled, staffing patterns, teaching vacancies, high school graduation rates, programs and services offered, and college application rates. Schools that are the only school in the district, state-run schools (e.g., schools for the blind), charter schools that do not report to a traditional school district, and BIE-funded schools received the School Questionnaire (with district items), an expanded version of the Public School Questionnaire that included items from the School District Questionnaire.

Teacher Survey. The questionnaire for this survey is mailed to a sample of teachers from the SASS sample of schools. It is sent to teachers in public schools, private schools, charter schools, and BIE schools. The Teacher Questionnaire collects data from teachers about their education and training, teaching assignment, certification, workload, and perceptions

and attitudes about teaching. Questions are also asked about teacher preparation, induction, organization of classes, computers, and professional development. The only eligible respondent for each teacher questionnaire is the teacher named on the questionnaire label. As of the 1993–94 SASS, administrators are eligible for both the Teacher Survey and the Principal Survey, if they teach a regularly scheduled class. In the 2007–08 Teacher Survey, items on grade range of teaching certification, use of electronic communications with parents, and out-of-pocket expenses for school supplies were added or revised.

Teacher Listing Form. The SASS Teacher Listing Form collects the full list of teachers from a school, along with information on subject matter taught, full- or part-time teaching status, and teaching experience. A question about teachers' race/ethnicity was replaced in the 2007–08 data collection by a question about teachers' status for the next school year. The information in the Teacher Listing Form is used to select a representative teacher sample and send out the Teacher Questionnaires. In 2007–08, the Teacher Listing Form restored a section that was removed in 2003–04, which had asked about the school name and grade range for verification purposes. (This section was not included in the survey questionnaire in 2003–04, as it was verified at the school, using a laptop-collected form.)

Additional Components

In addition to the core data collection described above, SASS featured additional components focusing on library media specialists/librarians and on student records in 1993–94 and on library media centers in 1993–94, 1999–2000, 2003–04, and 2007–08. One year following each SASS, a Teacher Follow-up Survey (TFS) is mailed to a sample of participants in the SASS Teacher Survey. (See chapter 5 for a complete description of TFS.) In 2007–08, SASS also included a Principal Follow-up Survey.

School Library Media Center Survey. This survey was added in the 1993–94 SASS. The questionnaire for the survey asks public and BIE schools about their access to and use of new information technologies. The questionnaire was not sent to private schools in 2003–04, due to budgetary reasons. (In 2007–08, the survey only surveyed public schools as well.) The survey collects data on library collections, media equipment, use of technology, staffing, student services, expenditures, currency of the library collection, and collaboration between the library media specialist and classroom teachers. A section on information literacy was added to the 2003–04 questionnaire. Items on access to online licensed databases, resource

availability, and information literacy were added or revised in the 2007–08 questionnaire. (See chapter 10 for a more complete description of this survey.)

School Library Media Specialist/Librarian Survey.

The questionnaire for this survey was mailed to a subsample of the SASS sample of public, private, and BIE schools in 1993–94. The survey solicited data that could be used to describe school librarians—for example, their educational background, work experience, and demographic characteristics. Because much of the collected information was comparable to that obtained in the Teacher Questionnaire, comparisons between librarians and classroom teachers can be made.

Periodicity

Between the 1987–88 and 1993–94 school years, SASS core components were on a 3-year cycle, with the TFS conducted 1 year after SASS. After a 6-year hiatus, SASS was fielded again in the 1999–2000, 2003–04, and 2007–08 school years (with the TFS following in 2000–01, 2004–05, and 2009–10). Subsequent SASS administrations are scheduled on a 4-year cycle.

2. USES OF DATA

SASS is the largest, most extensive survey of school districts, schools, principals, teachers, and library media centers in the United States today. It includes data from the public, private, and BIE school sectors. Moreover, SASS is the only survey that studies the complete universe of public charter schools. Therefore, SASS provides a multitude of opportunities for analysis and reporting on issues related to elementary and secondary schools.

SASS data have been collected six times between 1987 and 2007. Many questions have been asked of respondents at multiple time points, allowing researchers to examine trends on these topics over time. SASS asks similar questions of respondents across sectors, including public, public charter, BIE, and private schools. The consistency of questions across sectors and the large sample sizes allow for exploration of similarities and differences across sectors.

SASS data are representative at the state level for public school respondents and at the private school affiliation level for private school respondents. Thus, SASS is invaluable for analysts interested in elementary, middle, and secondary schools within or across specific states or private school affiliations. The

large SASS sample sizes allow extensive disaggregation of data according to the characteristics of teachers, administrators, schools, and school districts. For example, researchers can compare urban and rural settings and the working conditions of teachers and administrators of differing demographic backgrounds.

SASS collects extensive data on teachers, principals, schools, and school districts. Information on teachers includes their qualifications, early teaching experience, teaching assignments, professional development, and attitudes about the school. The SASS School Principal Questionnaire collects information about principals' or school heads' years of experience and training, goals and decision making, professional development for teachers and instructional aides, school climate and safety, student instructional time, principal perceptions and working conditions, and demographic information. Questions about schools include enrollment, staffing, the types of programs and services offered, school leadership, parental involvement, and school climate. At the district level, information is sought on the recruitment and hiring of teachers, professional development programs, student services, and other relevant topics.

SASS data can be very useful for researchers performing their own focused studies on smaller populations of teachers, administrators, schools, or school districts. SASS can supply data at the state, affiliation, or national level that provide valuable contextual information for localized studies; localized studies can provide illustrations of broad findings produced by SASS.

Users of restricted-use SASS data can link school districts and schools to other data sources. For instance, 2007–08 SASS restricted-use datasets include selected information taken from the CCD, but researchers can augment the datasets by adding more data from the CCD—either fiscal or nonfiscal data.

3. KEY CONCEPTS

Because of the large number of concepts in SASS surveys, only those pertaining to the level of data collection (LEA, school, teacher, library) are described in this section. For additional terms, the reader is referred to glossaries in SASS reports.

Local Education Agency (LEA). A public school district, or LEA, is defined as a government agency employing elementary- and secondary-level teachers

and administratively responsible for providing public elementary and/or secondary instruction and educational support services. Districts that do not operate schools but employ teachers were last included in the 1999–2000 SASS. (For example, some states have special education cooperatives that employ special education teachers who teach in schools in more than one school district.)

Public School. An institution that provides educational services for at least one of grades 1–12 (or comparable ungraded levels), has one or more teachers to give instruction, is located in one or more buildings, receives public funds as primary support, and is operated by an education agency. Schools in juvenile detention centers and schools located on military bases and operated by the Department of Defense are included.

Private School. An institution that is not in the public system and that provides instruction for any of grades 1–12 (or comparable ungraded levels). The instruction must be given in a building that is not used primarily as a private home. Private schools are divided into three categories: (1) Catholic: parochial, diocesan, private order; (2) other religious: affiliated with a conservative Christian school association, affiliated with a national denomination, unaffiliated; and (3) nonsectarian: regular, special program emphasis, special education. The classification of nonsectarian schools by program emphasis disentangles private schools offering a conventional academic program (regular) from those that either serve special-needs children (special education) or provide a program with a special emphasis (e.g., arts and sciences).

Charter School. A charter school is a public school that, in accordance with an enabling state statute, has been granted a charter exempting it from selected state or local rules and regulations. A charter school may be a newly created school or it may previously have been a public or private school.

BIE School. A school funded by the Bureau of Indian Education of the Bureau of Indian Affairs, U.S. Department of the Interior. These schools may be operated by the BIE, a tribe, a private contractor, or an LEA.

Library Media Center. A library media center is an organized collection of printed, audiovisual, or computer resources that (a) is administered as a unit, (b) is located in a designated place or places, and (c) makes resources and services available to students, teachers, and administrators.

Teacher. A full- or part-time teacher who teaches any regularly scheduled classes in any of grades K–12.¹ This includes administrators, librarians, and other professional or support staff who teach regularly scheduled classes on a part-time basis. Itinerant teachers are also included, as well as long-term substitutes who are filling the role of a regular teacher on a long-term basis. An itinerant teacher is one who teaches at more than one school (e.g., a music teacher who teaches 3 days per week at one school and 2 days per week at another). Short-term substitute teachers and student teachers are not included.

4. SURVEY DESIGN

Target Population

LEAs that employ elementary- and/or secondary-level teachers (e.g., public school districts, state agencies that operate schools for special student populations, such as inmates of juvenile correctional facilities or students in Department of Defense schools); cooperative agencies that provide special services to more than one school district; public, private, BIE, and charter schools with students in any of grades 1–12; the principals of these schools; library media centers; and teachers in public, private, BIE, and charter schools who teach students in grades K–12 in a school with at least a 1st grade.

Sample Design

SASS uses a stratified probability sample design. Details of stratification variables, sample selection, and frame sources are provided below.

Public school sample. In the public school sample, schools are selected first. The first level of stratification is by type of school: (a) BIE schools (all BIE schools are automatically in the sample); (b) schools with a high percentage of American Indian students (i.e., schools with 19.5 percent or more American Indian students); (c) schools in Delaware, Florida, Maryland, Nevada, and West Virginia (where it is necessary to implement a different sampling methodology to select at least one school from each LEA in the state); (d) charter schools; and (e) all other schools. Schools falling into more than one group are assigned to types A, B, D, C, and E in that order. The second level of stratification varies within school type. All BIE schools are automatically selected for the sample, so no stratification is needed. Schools with a high percentage of American Indian students are stratified by state

(Arizona; California; Montana; New Mexico; Washington; the remaining western states; Minnesota; North Dakota; South Dakota; the remaining midwestern states; North Carolina; Oklahoma; and the remaining states except Alaska, since most Alaskan schools have a high Native American enrollment). Schools in Delaware, Florida, Maryland, Nevada, and West Virginia are stratified first by state and then by LEA. Charter schools and schools not placed in another category are stratified by state. Within each second level, there are three grade level strata (elementary, secondary, and combined schools).

Within each stratum, all non-BIE schools are systematically selected using a probability proportionate to size algorithm. The measure of size used for schools in the CCD is the square root of the number of teachers in the school as reported in the CCD file. Any school with a measure of size larger than the sampling interval is excluded from the probability sampling operation and included in the sample with certainty.

The CCD Public Elementary/Secondary School Universe Survey serves as the public school sampling frame. (See chapter 2 for a complete description of the CCD.) The frame includes regular public schools, Department of Defense-operated military base schools, and special purpose schools (such as special education, vocational, and alternative schools). Schools outside the United States and schools that teach only prekindergarten, kindergarten, or postsecondary students are deleted from the file. The following years of the CCD were used as the public school frame for the last five rounds of SASS:

- 2005–06 CCD for the 2007–08 SASS;
- 2001–02 CCD for the 2003–04 SASS;
- 1997–98 CCD for the 1999–2000 SASS;
- 1991–92 CCD for the 1993–94 SASS; and
- 1988–89 CCD for the 1990–91 SASS.

In the 1987–88 SASS, the 1986 Quality Education Data (QED) survey was used as the sampling frame (Kaufman 1991).

Private school sample. For private schools, the sample is stratified within each of the two types of frames: (1) a list frame, which is the primary private school frame; and (2) an area frame, which is used to identify schools not included in the list frame and to compensate for the undercoverage of the list frame. Private schools in the

¹ A teacher teaching only kindergarten students is in scope, provided the school serves students in a grade higher than kindergarten.

list frame are stratified by affiliation, grade level, and region. Within each stratum, schools are sampled systematically using a probability proportionate to size algorithm. Any school with a measure of size larger than the sampling interval is excluded from the probability sampling process and included in the sample with certainty. All schools in the area frame within noncertainty PSUs and not already listed in the list frame are included in the sample with certainty.

The most recent PSS, updated with the most recent association lists, serves as the private school sampling frame. For example, the 2001–02 PSS—updated with 26 lists of private schools provided by a private school association (as well as 51 lists of private schools, from the 50 states and the District of Columbia)—was used as the private school frame for the 2003–04 SASS. For the 2007–08 SASS, the private school list frame was based on the 2005–06 PSS, updated with private school organizations and state lists collected by the U.S. Census Bureau in the summer of 2006. The 1991–92, 1989–90, and 1997–98 PSS were the basis for the private school frame for the 1993–94, 1990–91, and 1999–2000 SASS, respectively. The 1986 QED survey was used as the sampling frame for the 1987–88 SASS.

BIE school selection. Since the 1993–94 SASS, all BIE schools have been selected with certainty; in 1990–91, 80 percent of BIE schools were sampled. The BIE school frame for the 2003–04 SASS consisted of a list of schools that the BIE operated or funded during the 2001–02 school year. (The list was obtained from the U.S. Department of the Interior.) The BIE list was matched against the CCD, and the schools on the BIE list that did not match the CCD were added to the universe of schools.

For the 2007–08 SASS data collection, a separate universe of schools operated or funded by the BIE in the 2005–06 school year was drawn from the Program Education Directory maintained by the BIE. (The CCD now defines the BIE as its own “territory,” similar to Puerto Rico and other non-state territories, and does not permit duplicates to be reported by the states.) All BIE schools meeting the SASS definition of a school were included in the sample.

Charter school selection. In the 1999–2000 SASS, a charter school sample was added. All charter schools were selected with certainty from the frame, which consisted of a list of charter schools developed for the U.S. Department of Education’s Institute of Education Sciences. The list included only charter schools that were open (teaching students) during the 1998–99 year. This changed in the 2003–04 SASS, when a nationally representative sample of public charter schools was

included as part of the public school sample. In the 2007–08 SASS, charter schools continued to be included as a part of the public school sample.

Each school sampled for SASS receives a school questionnaire, and the principal of each sampled school receives a principal questionnaire.

Teacher selection. Within each sampled school, a sample of teachers is selected. First, the sampled schools are asked to provide a list of their teachers and selected characteristics. For example, in the 2007–08 SASS data collection, the Teacher Listing Form was collected as early as possible in the 2007–08 school year at all public (including public charter), private, and BIE-funded schools in the SASS sample to obtain a complete list of all the teachers employed at each school.

In the 2007–08 SASS, teachers were stratified into one of two teacher types: new and experienced. For new and experienced teachers in public schools, oversampling was not required, due to the large number of sampled schools with new teachers. Therefore, teachers were allocated to the new and experienced categories in proportion to their numbers in the school. However, in private schools, new teachers were oversampled. Before teachers were allocated to the new or experienced strata, schools were first allocated an overall number of teachers to be selected.

Teacher records within a school are sorted by the teacher stratum code, the teacher subject code, and the teacher line number code. The teacher line number code is a unique number assigned to identify the teacher within the list of teachers keyed by the field representative. Within each teacher stratum in each school, teachers are selected systematically with equal probability. The within-school probabilities of selection are computed so as to give all teachers within a school stratum the same overall probability of selection (self-weighted) within teacher and school strata, but not across strata. However, since the school sample size of teachers is altered due to the minimum constraint (i.e., at least one teacher per school) or maximum constraint (i.e., no more than either twice the average stratum allocation or 20 teachers per school), the goal of achieving self-weighting for teachers is lost in some schools. Each sampled teacher receives a teacher questionnaire.

Library media center selection. For the 2003–04 and 2007–08 SASS, all library media centers in public, public charter, and BIE-funded schools in the SASS

sample were asked to complete the School Library Media Center Questionnaire.

School district selection. In most states, once public schools are selected, the districts associated with these schools are placed in the sample as well. However, in Delaware, Florida, Maryland, Nevada, and West Virginia, all districts are defined as school sampling strata, placing all districts in each of these states in the district sample. (In some SASS administrations, a sample of districts not associated with schools is taken, but not in the 2007–08 SASS.) The district sample is selected using a probability proportionate to size algorithm. Each sampled school district receives a school district questionnaire.

The approximate sample sizes for the 2007–08 SASS were 12,900 schools and administrators, some 56,370 teachers, and 5,250 school districts.

Data Collection and Processing

The 2007–08 SASS was primarily a mailout/mailback survey with computer-assisted telephone interviewing (CATI) and telephone follow-up. In 2003–04 and 2007–08, the School Library Media Center Survey did not have an Internet reporting option, as it did in 1999–2000. All survey modes used in SASS are administered by the U.S. Bureau of the Census.

Reference Dates. Data for SASS components are collected during a single school year. Most data items refer to that school year. Questions on enrollment and staffing refer to October 1 of the school year. Questions for teachers about current teaching loads refer to the most recent full week that school was in session, and questions on professional development refer to the past 12 months.

Data Collection. The data collection procedures begin with advance mailings to school districts explaining the nature and purpose of SASS. Field staff then phone school principals to set up face-to-face appointments with them. The telephone call includes a request to prepare a list of all eligible teachers in their schools. If the teacher roster is not provided at the appointment, field staff make arrangements to obtain the roster at a later meeting. The teacher sample is selected using these lists.

The school district questionnaires are mailed out first. Then, the school, principal, and library media center surveys are delivered to schools in person. The teacher questionnaires are delivered last. Follow-up efforts begin approximately 2 weeks after questionnaires are distributed. They consist of telephone calls and personal visits to schools to obtain completed

questionnaires or to verify that they have been mailed back. Field staff record the status of each questionnaire and, if necessary, supply additional blank questionnaires.

Processing. During the check-in phase, each questionnaire is assigned an outcome code: completed interview, out-of-scope, or noninterview. A combination of manual data keying and imaging technology was used to enter the data. Then, interview records in the data files undergo a round of primary data review, where analysts examine the frequencies of each data item in order to identify any suspicious values. Census staff review the problem cases and make corrections whenever possible.

After the primary data review, all records (i.e., records from all survey components) classified as interviews are subject to a set of computer edits: a range check, a consistency edit, and a blanking edit. After the completion of these edits, the records are put through another edit to make a final determination of whether the case is eligible for the survey, and, if so, whether sufficient data have been collected for the case to be classified as an interview. A final interview status recode (ISR) value is assigned to each case as a result of the edit.

Estimation Methods

Sample units are weighted to produce national and state estimates for public elementary and secondary school surveys (i.e., schools, teachers, administrators, school districts, and school library media centers); and national estimates for BIE, charter school, and public combined school surveys (i.e., schools, teachers, administrators, and school library media centers). The private sector is weighted to produce national and affiliation group estimates. These estimates are produced through the weighting and imputation procedures discussed below.

Weighting. Estimates from SASS sample data are produced by using weights. The weighting process for each component of SASS includes adjustments for nonresponse using respondents' data and adjustments of the sample totals to the frame totals to reduce sampling variability. The exact formula representing the construction of the weight for each component of SASS is provided in each administration's sample design report (e.g., *1993–94 Schools and Staffing Survey: Sample Design and Estimation* [Abramson et al. 1996]). The construction of weights is also discussed in the *Quality Profile* reports (Jabine 1994; Kalton et al. 2000) and in the documentation for the 2003–04 administration (Tourkin et al. 2007). Since SASS and PSS data were collected at the same time in

1993–94 and 1999–2000, in both years the number of private schools reported in SASS was made to match the number of private schools reported in PSS.

Imputation. In all administrations of SASS, all items with missing values are imputed for records classified as interviews. SASS uses a two-stage imputation procedure. The first-stage imputation uses a logical or deductive method, such as:

- Using data from other items in the same questionnaire;
- Extracting data from a related SASS component (different questionnaire); or
- Extracting information about the sample case from the PSS or CCD, the sampling frames for private and public schools, respectively.

In addition, some inconsistencies between items are corrected by ratio adjustment during the first-stage imputation.

The second-stage imputation process is applied to all items with missing values that were not imputed in the first stage. This imputation uses a hot-deck imputation method, extracting data from a respondent (i.e., a donor) with similar characteristics to the nonrespondent. If there is still no observed value after collapsing to a certain point, the missing values are imputed using a clerically imputed value or automated algorithm.

Recent Changes

Several changes were made over time, largely due to budgetary reasons.

Design changes from 1999–2000 to 2007–08:

- Rather than surveying all public charter schools, as was done in the 1999–2000 SASS, some 300 public charter schools were sampled for the 2003–04 SASS.
- The separate questionnaire for public charter schools was discontinued. The reduction in the public charter school sample size from 1,100 in the 1999–2000 SASS to about 300 in the 2003–04 SASS meant it was no longer feasible to produce a separate questionnaire, since public charter school data could not be published with as much detail (for the 2003–04 SASS, only at the national and regional levels). Public charter school data are now included with traditional public school data.

- Affiliation for private schools was redefined and stratified into 17 groups rather than the previous 20 groups in the 2003–04 SASS. Catholic schools were split into three groups based on typology. Other religious schools were divided into five groups corresponding to the four largest non-Catholic religious organizations (by number of schools) and a catch-all “other.” Nonsectarian schools were divided into three groups by typology.
- Grade-level stratification in public and private schools was defined purely on the basis of grade level of the school starting in 2003–04 SASS. Schools classified as a type other than “regular school” were no longer placed by default in the combined school category, which includes schools with some elementary and some secondary grades. Many nonregular schools (i.e., special education, alternative, and vocational schools) cover a specific grade range. To the extent this grade range is known, this was a more appropriate method of stratification than placing them all in the combined school strata. Nonregular schools with a grade range that is ungraded or unknown remain in the combined school strata.
- Public schools from the CCD were collapsed into what was perceived to be a better fit with the SASS definition of a school prior to stratification beginning in the 2003–04 SASS. The sample allocation was revised to avoid undersampling schools now classified at the combined grade level. In other words, the revision of the sample allocation ensured that the newly combined schools were sampled at the same approximate rate as they would have been prior to the collapsing procedure. In general, the combined school sample size was increased to the point at which the combined school sampling rate equaled the overall state-level sampling rate. For example, if one in five schools were sampled in a particular state, then one in five of the combined schools were sampled rather than using the default sample size of 10 combined schools.
- The sort order for the public and private school sampling was altered to sort on enrollment in a serpentine fashion (instead of always sorting in descending order) in the 2003–04 SASS. Serpentine sorting involves sorting in ascending order with respect to higher level sort variables one time, then sorting in descending order the next time, and so on. This

reduces the variation in enrollment between adjacent sampled schools and thus reduces the overall sampling error.

- Florida and Maryland were added to the list of states where at least one school is selected in each school district. This was done in the 2003–04 SASS to decrease the standard error of the state-level school district estimates.
- Oversampling of bilingual/English as a Second Language (ESL) teachers was discontinued in the 2003–04 SASS, since a sufficient number of bilingual teachers to produce the desired reliability estimates could be done without oversampling.
- Teacher sampling was automated to speed up the distribution of the teacher questionnaires. This, however, reduced the level of control over the sample sizes for the remaining oversampled teacher strata (Asian/Pacific Islander and American Indian/Alaska Native). The automation no longer allowed the sampling rate for these teachers to be periodically revised during the sampling process. Thus, if the number of these teachers listed differed from the expected number, the sample size goal would no longer be met.
- The School Library Media Center Questionnaire was not administered to private schools for budget reasons as of the 2003–04 SASS.
- The School Questionnaire (with district items) is a questionnaire that contains the public school questions and most of the school district questions in the 1999–2000 SASS. It was administered to public charter, state-operated (often schools for the blind or schools located in juvenile detention facilities), and BIE-funded schools, as well as public schools in one-school districts. This change was made to ease respondent burden in cases where the respondent for the school and school district questionnaires was expected to be the same.

Future Plans

SASS administrations are now scheduled on a 4-year cycle. The next administration will be in 2011–12.

5. DATA QUALITY AND COMPARABILITY

Sampling Error

The estimators of sampling variances for SASS statistics take the SASS complex sample design into account. For an overview of the calculation of sampling errors, see the *Quality Profile* reports (Jabine 1994; Kalton et al. 2000).

Direct Variance Estimators. The balanced half-sample replication (BHR) method, also called balanced repeated replication (BRR), was used to estimate the sampling errors associated with estimates from the 1987–88 and 1990–91 SASS. Given the replicate weights, the statistic of interest (e.g., the number of 12th grade teachers from the School Survey) can be estimated from the full sample and from each replicate. The mean square error of the replicate estimates around the full sample estimate provides an estimate of the variance of the statistic.

A bootstrap variance estimator was used for the 1993–94, 1999–2000, 2003–04, and 2007–08 SASS. The bootstrap variance reflects the increase in precision due to large sampling rates because the bootstrap is done systematically without replacement, as was the original sampling. Bootstrap samples can be selected from the bootstrap frame, replicate weights computed, and variances estimated with standard BHR software. The bootstrap replicate basic weights (inverse of the probability of selection) were subsequently reweighted. More information on the bootstrap variance methodology and how it applies to SASS is contained in the following sources: “A Bootstrap Variance Estimator for Systematic PPS Sampling” (U.S. Department of Education 2000) which describes the methodology used in the 1999–2000 SASS; “A Bootstrap Variance Estimator for the Schools and Staffing Survey” (U.S. Department of Education 1994); “Balanced Half-Sample Replication With Aggregation Units” (U.S. Department of Education 1994); “Comparing Three Bootstrap Methods for Survey Data” (Sitter 1990); “Properties of the Schools and Staffing Survey Bootstrap Variance Estimator” (U.S. Department of Education 1996); and “The Jackknife, the Bootstrap and Other Resampling Plans” (Efron 1982).

SASS variances can be calculated using the replicates of the full sample that are available in the data files with software such as WesVarPC. For examples of other software that support BRR, see *Introduction to Variance Estimation* (Wolter 1985).

Average Design Effects. Design effects (*Deffs*) measure the impact of the complex sample design on the accuracy of a sample estimate, in comparison to the alternative simple random sample design. For the 1990–91 SASS, an average design effect was derived for groups of statistics and, within each group, for a set of subpopulations. Standard errors for 1990–91 and 1993–94 SASS statistics of various groups for various subpopulations can then be calculated approximately from the standard errors based on the simple random sample (using SAS or SPSS) in conjunction with the average design effects provided. For example, for the 1990–91 SASS, average design effects for selected variables in the School Survey are 1.60 (public sector) and 1.36 (private sector); in the Principal Survey, 4.40 (public sector) and 4.02 (private sector); and in the Teacher Survey, 3.75 (public sector) and 2.52 (private sector). Examples illustrating the use of SASS average design effect tables are provided in *Design Effects and Generalized Variance Functions for the 1990–91 Schools and Staffing Survey (SASS), Volume I, User’s Manual* (Salvucci and Weng 1995).

Generalized Variance Functions (GVFs). GVF tables were developed for use in the calculation of standard errors of totals, averages, and proportions of interest in the 1990–91 SASS components. The 1990–91 GVFs can be used for the 1993–94 SASS because no major design changes were adopted between 1990–91 and 1993–94. Note that the GVF approach, unlike the design effect approach described above, involves no need to calculate the simple random sample variance estimates. Examples illustrating the use of the GVF tables are provided in *Design Effects and Generalized Variance Functions for the 1990–91 Schools and Staffing Survey (SASS), Volume I, User’s Manual* (Salvucci and Weng 1995).

Nonsampling Error

Coverage Error. SASS surveys are subject to any coverage error present in the CCD and PSS data files, which serve as their principal sampling frames. The report *Coverage Evaluation of the 1994–95 Common Core of Data: Public Elementary/Secondary Education Agency Universe Survey* (Owens 1997) found that overall coverage in the 1994–95 CCD Local Education Agency Universe Survey was 96.2 percent (in a comparison to state education directories). “Regular” agencies—those traditionally responsible for providing public education—had almost total coverage in the 1994–95 agency universe survey. Most coverage discrepancies were attributed to nontraditional agencies that provide special education, vocational education, and other services. However, there is potential for undercoverage bias associated with the absence of schools built between the time when the sampling

frame is constructed and the time of the SASS survey administration. Further research on coverage can be found in “Evaluating the Coverage of the U.S. National Center for Education Statistics’ Public Elementary/Secondary School Frame” (Hamann 2000) and “Evaluating the Coverage of the U.S. National Center for Education Statistics’ Public and Private School Frames Using Data from the National Assessment of Educational Progress” (Lee, Burke, and Rust 2000).

A capture-recapture methodology was used to estimate the number of private schools in the United States and to estimate the coverage of private schools in the 1999–2000 PSS; the study found that the PSS school coverage rate is equal to 97 percent. (See chapter 2 for a description of the CCD and chapter 3 for a description of the PSS.)

Nonresponse Error.

Unit nonresponse. The weighted unit response rates for public schools have been higher than the weighted unit response rates for private schools in all six rounds of SASS. (See table 3 for response rates from selected years.) For more information on the analysis of nonresponse rates, refer to *An Analysis of Total Nonresponse in the 1993–94 Schools and Staffing Survey (SASS)* (Monaco et al. 1997) and *An Exploratory Analysis of Response Rates in the 1990–91 Schools and Staffing Survey (SASS)* (Scheuren et al. 1996).

Item Nonresponse. For the 2007–08 SASS, the weighted item response rates for the individual surveys were as follows: 52 to 100 percent for public school districts; 71 to 100 percent for public schools; 49 to 100 percent for private schools; 65 to 100 percent for BIE schools; 76 to 100 percent for public school principals; 86 to 100 percent for private school principals; and 61 to 100 percent for BIE school principals. For teachers, the ranges of item response rates were as follows: 44 to 100 percent for public school teachers; 64 to 100 percent for private school teachers; and 0 to 100 percent for BIE teachers. Item response rates for public school library media centers and BIE school library media centers ranged from 84 to 100 percent and 71 to 100 percent, respectively.

Measurement Error. Results reported in *An Analysis of Total Response in the 1993–94 Schools and Staffing Survey (SASS)* (Monaco et al. 1997) support the contention that, without follow-up to mail surveys, nonresponse error would be much greater than it is and that the validity and reliability of the data would be considerably reduced. However, because of the substantial amount of telephone follow-up, there is

Table 3. Summary of weighted response rates for selected SASS questionnaires

Questionnaire	1993 -94	1999 -2000	2003 -04	2007 -08
School District Survey	93.9	88.6	82.9	87.8
Public Principal Survey	96.6	90.0	82.2	79.4
Public School Survey	92.3	88.5	80.8	80.4
Public Teacher Survey ¹	83.8	83.1	75.7	84.0
Private Principal Survey	87.6	84.8	74.9	72.2
Private School Survey	83.2	79.8	75.9	75.9
Private Teacher Survey ¹	72.9	77.2	70.4	77.5
BIE Principal Survey	98.7	93.3	90.7	79.2
BIE School Survey	99.3	96.7	89.5	77.1
BIE Teacher Survey	86.5	87.4	86.3	81.8

¹The overall teacher response rates are the percentage of teachers responding in schools that provided teacher lists for sampling. SOURCE: Aritomi, P., and Coopersmith, J. (2009). *Characteristics of Public School Districts in the United States: Results From the 2007-08 Schools and Staffing Survey* (NCES 2009-320). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Gruber, K.J., Rohr, C.L., and Fondelier, S.E. (1996). *1993-94 Schools and Staffing Survey: Data File User's Manual* (NCES 96-142). National Center for Education Statistics, U.S. Department of Education. Washington, DC; Tourkin, S.C., Pugh, K.W., Fondelier, S.E., Parmer, R.J., Cole, C., Jackson, B., Warner, T., and Weant, G. (2004). *1999-2000 Schools and Staffing Survey (SASS) Data File User's Manual* (NCES 2004-303). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Tourkin, S.C., Warner, T., Parmer, R., Cole, C., Jackson, B., Zukerberg, A., Cox, S., and Soderborg, A. (2007). *Documentation for the 2003-04 Schools and Staffing Survey* (NCES 2007-337). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

concern about possible bias due to differences in the mode of survey collection. Other possible sources of

measurement error include long, complex instructions that respondents either do not read or do not understand, navigation problems related to the format of the questionnaires, and definitional and classification problems. See also *Measurement Error Studies at the National Center for Education Statistics* (Salvucci et al. 1997).

Several NCES working papers also address measurement error. Reports on the 1993-94 SASS include *Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey* (Jenkins and Von Thurn 1996); *Further Cognitive Research on the Schools and Staffing Survey (SASS)* (Zukerberg and Lee 1997); *Report of Cognitive Research on the Public and Private School Teacher Questionnaires for the Schools and Staffing Survey 1993-94 School Year* (Jenkins 1997); and *Response Variance in the 1993-94 Schools and Staffing Survey: A Reinterview Report* (Bushery, Schreiner, and Sebron 1998). Reports on the 1991-92 SASS include the *1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report* (Royce 1994) and *The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation* (Jenkins and Wetzel 1995).

6. CONTACT INFORMATION

For content information on SASS, contact:

Kerry Gruber
Phone: (202) 502-7349
E-mail: kerry.gruber@ed.gov
SASS e-mail: sassdata@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Broene, P. (1991a). *User's Manual: 1987-88 Schools and Staffing Survey - Public School Administrator Questionnaire Data* (NCES 91-137). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

- Broene, P. (1991b.) *User's Manual: 1987–88 Schools and Staffing Survey - Public School Data* (NCES 91-136). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Gruber, K.J., Rohr, C.L., and Fondelier, S.E. (1994). *1990–91 Schools and Staffing Survey: Data File User's Manual Volume I: Survey Documentation* (NCES 93-144-I). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Gruber, K.J., Rohr, C.L., and Fondelier, S.E. (1996). *1993–94 Schools and Staffing Survey: Data File User's Manual, Volume I: Survey Documentation* (NCES 96-142). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Rouk, Ü., Weiner, L., and Riley, D. (1999). *What Users Say About Schools and Staffing Survey Publications* (NCES Working Paper 1999-10). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Tourkin, S.C., Pugh, K.W., Fondelier, S.E., Parmer, R.J., Cole, C., Jackson, B., Warner, T., and Weant, G. (2004). *1999–2000 Schools and Staffing Survey (SASS) Data File User's Manual* (NCES 2004-303). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Tourkin, S., Thomas, T., Swaim, N., Cox, S., Parmer, R., Jackson, B., Cole, C., and Zhang, B. (2010). *Documentation for the 2007-08 Schools and Staffing Survey* (NCES 2010-332). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Tourkin, S.C., Warner, T., Parmer, R., Cole, C., Jackson, B., Zukerburg, A., Cox, S., and Soderborg, A. (2007). *Documentation for the 2003–04 Schools and Staffing Survey* (NCES 2007-337). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Uses of Data**
- Ingersoll, R.M. (1998). *An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey* (NCES Working Paper 98-18). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- McGrath, D.J., and Luekens, M.T. (2000). *A Research Agenda for the 1999–2000 Schools and Staffing Survey* (NCES Working Paper 2000-10). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Mullens, J.E., and Kasprzyk, D. (1997). *The Schools and Staffing Survey: Recommendations for the Future* (NCES 97-596). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Wiley, S.D., and Reynolds, K.A. (1999). *Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results* (NCES Working Paper 1999-02). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Survey Design**
- Abramson, R., Cole, C., Fondelier, S., Jackson, B., Parmer, R., and Kaufman, S. (1996). *1993–94 Schools and Staffing Survey: Sample Design and Estimation* (NCES 96-089). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Baker, D.P., Levine, R., Han, M., and Garet, M. (1998). *A Feasibility Study of Longitudinal Design for Schools and Staffing Survey* (NCES Working Paper 98-16). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Ingersoll, R.M. (1995). *An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey* (NCES Working Paper 95-18). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Ingersoll, R.M. (1996). *National Assessments of Teacher Quality* (NCES Working Paper 96-24). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Isaacs, J.B., Best, C.M., Cullen, A.D., Garet, M.S., and Sherman, J.D. (1998). *Collection of Public School Expenditure Data: Development of a Questionnaire* (NCES Working Paper 98-01). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Isaacs, J.B., Garet, M.S., Sherman, J.D., Cullen, A., and Phelps, R. (1999). *Collection of Resource and Expenditure Data on the Schools and Staffing Survey* (NCES Working Paper 1999-07). National

Center for Education Statistics, U.S. Department of Education. Washington, DC.

Kaufman, S. (1991). *1988 Schools and Staffing Survey Sample Design and Estimation* (NCES 91-127). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Kaufman, S., and Huang, H. (1993). *1990–91 Schools and Staffing Survey: Sample Design and Estimation* (NCES 93-449). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Levine, R.E., Chambers, J.G., Dueñas, I.E., and Hikido, C.S. (1997). *Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey* (NCES Working Paper 97-42). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Rollefson, M. (1998). *The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper* (NCES Working Paper 98-08). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Bushery, J., Schreiner, I., and Sebron, G. (1998). *Response Variance in the 1993–94 Schools and Staffing Survey: A Reinterview Report* (NCES Working Paper 98-02). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Efron, B. (1982). The Jackknife, the Bootstrap, and Other Resampling Plans. *Society of Industrial and Applied Mathematics CBMS-NSF Monographs*, 38.

Hamann, T.A. (2000). Evaluating the Coverage of the U.S. National Center for Education Statistics' Public Elementary/Secondary School Frame. In *The Second International Conference on Establishment Surveys: Survey Methods for Businesses, Farms, and Institutions* (pp. 79–88). Arlington, VA: American Statistical Association.

Jabine, T. (1994). *Quality Profile for SASS: Aspects of the Quality of Data in the Schools and Staffing Surveys (SASS)* (NCES 94-340). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Jenkins, C.R. (1997). *Report of Cognitive Research on the Public and Private School Teacher Questionnaires for the Schools and Staffing Survey 1993–94 School Year* (NCES Working Paper 97-10). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Jenkins, C.R., and Von Thurn, D. (1996). *Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey* (NCES Working Paper 96-05). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Jenkins, C.R., and Wetzel, A. (1995). *The Results of the 1991–92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation* (NCES Working Paper 95-10). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Kalton, G., Winglee, M., Krawchuk, S., and Levine, D. (2000). *Quality Profile for SASS Rounds 1–3: 1987–1995, Aspects of the Quality of Data in the Schools and Staffing Surveys (SASS)* (NCES 2000-308). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Lee, H., Burke, J., and Rust, K. (2000). Evaluating the Coverage of the U.S. National Center for Education Statistics' Public and Private School Frames Using Data From the National Assessment of Educational Progress. In *The Second International Conference on Establishment Surveys: Survey Methods for Businesses, Farms, and Institutions* (pp. 89–98). Arlington, VA: American Statistical Association.

Monaco, D., Salvucci, S., Zhang, F., and Hu, M. (1997). *An Analysis of Total Nonresponse in the 1993–94 Schools and Staffing Survey (SASS)* (NCES 98-243). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Owens, S. (1997). *Coverage Evaluation of the 1994–95 Common Core of Data: Public Elementary/Secondary Education Agency Universe Survey* (NCES 97-505). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Royce, D. (1994). *1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report* (NCES 94-03). National Center for Education

- Statistics, U.S. Department of Education. Washington, DC.
- Salvucci, S., Conley, V., Fink, S., and Saba, M. (1997). *Measurement Error Studies at the National Center for Education Statistics* (NCES 97-464). National Center for Education Statistics, Sciences, U.S. Department of Education. Washington, DC.
- Salvucci, S., Holt, A., and Moonesingle, R. (1995). *Design Effects and Generalized Variance Functions for the 1990–91 Schools and Staffing Survey (SASS): Volume II, Technical Report* (NCES 95-342-II). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Salvucci, S., and Weng, S. (1995). *Design Effects and Generalized Variance Functions for the 1990–91 Schools and Staffing Surveys (SASS) Volume I—User’s Manual* (NCES 95-342-I). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Scheuren, F., Monaco, D., Zhang, F., Ikosi, G., and Chang, M. (1996). *An Exploratory Analysis of Response Rates in the 1990–91 Schools and Staffing Survey (SASS)* (NCES 96-338). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Sitter, R. (1990). Comparing Three Bootstrap Methods for Survey Data. *Technical Report Series of the Laboratory for Research in Statistics and Probability*. Carleton University.
- U.S. Department of Education, National Center for Education Statistics. (1994). A Bootstrap Variance Estimator for the Schools and Staffing Survey; Balanced Half-Sample Replication With Aggregation Units. In *Schools and Staffing Survey (SASS), Papers Presented at the Meetings of the American Statistical Association* (NCES Working Paper 94-01). Washington, DC: Author.
- U.S. Department of Education, National Center for Education Statistics. (1996). Properties of the Schools and Staffing Survey Bootstrap Variance Estimator. In *Schools and Staffing Survey (SASS): 1995 Selected Papers Presented at the 1995 Meeting of the American Statistical Association* (NCES Working Paper 96-02). Washington, DC: Author.
- U.S. Department of Education, National Center for Education Statistics. (2000). A Bootstrap Variance Estimator for Systematic PPS Sampling. In *Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings* (NCES Working Paper 2000-04). Washington, DC: Author.
- Wolter, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.
- Zukerberg, A., and Lee, M. (1997). *Further Cognitive Research on the Schools and Staffing Survey (SASS) Teacher Listing Form* (NCES Working Paper 97-23). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Chapter 5: SASS Teacher Follow-up Survey (TFS)

1. OVERVIEW

The SASS Teacher Follow-up Survey (TFS) is a follow-up survey of elementary and secondary school teachers who participated in the Schools and Staffing Survey (SASS) (see chapter 4 for details on SASS). TFS is conducted for the National Center for Education Statistics (NCES) by the U.S. Census Bureau in the school year following the SASS data collection. TFS consists of a subsample of teachers who left teaching within the year after the SASS was administered and a subsample of those who continued teaching, including those who remained in the same school as in the previous year and those who changed schools.

Purpose

To measure the attrition rate for teachers, examine the characteristics of teachers who stay in the teaching profession and those who leave, obtain activity or occupational data for those who leave the position of a K–12 teacher, obtain current teaching assignment information for those who are still teaching, and collect data on attitudes about the teaching profession in general and job satisfaction in particular. TFS is designed to support estimates of public elementary, secondary, and combined school teachers and private school teachers at the national level.

Components

TFS is composed of two questionnaires: the Former Teacher Questionnaire, which collects information from sampled teachers who leave the K–12 teaching profession within the year after SASS; and the Current Teacher Questionnaire, which collects information from sampled teachers who currently teach students in any of grades prekindergarten through 12. Eligible survey respondents are teachers in public and private elementary and secondary schools in the 50 states and the District of Columbia.

Former Teacher Questionnaire. This questionnaire collects information from former teachers on their current occupation, primary activity, plans to remain in their current position, plans for further education, plans for returning to teaching, reasons for leaving teaching, possible areas of satisfaction or dissatisfaction with teaching, salary, marital status, number of children, and reasons for retirement, as well as any other information that may be related to attrition.

Current Teacher Questionnaire. This questionnaire obtains information from current teachers, including teachers who continued to teach in the same school as in the previous year and those who changed schools. It collects information on occupational status (full time, part time), primary teaching assignment by field, teaching certificate, level of students taught, areas of satisfaction or dissatisfaction, new degrees earned or pursued, expected duration in teaching, marital status, number of children, academic year base salary, time spent performing school related tasks, and effectiveness of the school administration. If the teacher is teaching in a

TEACHER FOLLOW-UP SURVEY OF SCHOOL TEACHERS

TFS collects data on:

- Stayers
- Movers
- Leavers

different school than during the SASS administration, the questionnaire obtains information on the teacher's reasons for leaving the previous school.

Periodicity

TFS is a follow-up of selected teachers from the SASS teacher surveys and is conducted during the school year following the SASS administration. It was conducted in the 1988–89, 1991–92, 1994–95, 2000–01, and 2004–05 school years (after the 1987–88, 1990–91, 1993–94, 1999–2000, and 2003–04 administrations of SASS, respectively). The most recent survey was conducted in the 2008–09 school year, collecting data from a subsample of teachers who participated in the 2007–08 SASS.

2. USES OF DATA

Data from TFS are used for a variety of purposes by Congress, state education departments, federal agencies, private school associations, teacher associations, and educational organizations. TFS can be used to research issues related to teacher turnover. Leavers, movers, and stayers can be profiled and compared in terms of teaching qualifications, working conditions, attitudes toward teaching, job satisfaction, salaries, benefits, and other incentives and disincentives for remaining in or leaving the teaching profession. TFS also provides a measure of national teacher attrition in the various fields and updates information on the education, other training, and career paths of teachers. In addition, sampled teachers can be linked to SASS data to determine relationships between local district and school policies and practices, teacher characteristics, and teacher attrition and retention.

3. KEY CONCEPTS

Key Terms

Some of the key terms used in TFS are described below. For descriptions of other terms, see “Appendix A. Key Terms for TFS” in *Documentation for the 2008–09 Teacher Follow-up Survey* (forthcoming).

Leavers. Teachers who left the teaching profession or teachers who were no longer teaching in any of grades pre-K–12 in the school year after the SASS administration (includes teachers whose status

changed to short-term substitute, student teacher, or teacher aide).

Movers. Teachers who were still teaching in the school year after the SASS administration, but had moved to a different school.

Stayers. Teachers who were teaching in the same school in the year after the SASS administration as in the year of the SASS administration.

Itinerant teacher. An individual who teaches at more than one school; for example, a music teacher who teaches 3 days per week at one school and 2 days per week at another.

4. SURVEY DESIGN

Target Population

The target population is the universe of elementary and secondary school teachers who teach in public and private schools in the 50 states and the District of Columbia, in schools that had any of grades K–12 during the school year of the last SASS administration. This population is divided into two components: those who left teaching after that school year (former teachers) and those who continued teaching (current teachers).

The TFS sample of teachers includes those who left the position of a K–12 teacher in the year after SASS (leavers). It also includes those who continued to teach students in any of grades pre-K–12 or in comparable ungraded levels, including teachers who remained in the same school as in the previous year (stayers) and those who changed schools (movers). Prekindergarten is included so that sampled teachers who change assignments from teaching students in any of grades K–12 to teaching only prekindergarten students would not be considered leavers.

In SASS, the sampling frame for public schools is an adjusted version of the NCES Common Core of Data (CCD), and the sampling frame for private schools is a modified version of the NCES Private School Universe Survey (PSS). The sampling frame for the SASS teacher questionnaire consists of lists of teachers provided by schools in the SASS sample. A teacher is defined as a staff member who taught a regularly scheduled class to students in any of grades K–12 or comparable ungraded levels.

Sample Design

TFS surveys a sample of teachers who completed interviews in the previous year's SASS. The TFS sample is a stratified sample that is allocated to allow comparisons of teachers by five variables: status (stayers, movers, leavers, and unknown); school type (traditional public, public charter, and private); experience (new and experienced); grade level (elementary, middle, and secondary); and race/ethnicity (White, non-Hispanic, Black, Hispanic, and all other races/ethnicities). In the 2008–09 TFS administration, all responding SASS teachers in public schools who indicated that their first year of teaching was 2007 or 2008 were included in the sample. All other SASS responding teachers were stratified by the five variables in the following order: school type, teacher status, experience, teacher's grade level, and race/ethnicity.

Within each TFS stratum, teachers with completed interviews in SASS are sorted by a measure of size (the SASS teacher initial basic weight, which is the inverse of the probability of selection prior to any corrections identified during data collection), main subject taught as reported by the teacher in SASS (i.e., special education, general elementary, mathematics, science, English/language arts, social studies, vocational/technical, and other), Census region, SASS private school affiliation stratum (for private school teachers only), school locale (based on the 1990 Census geography), school enrollment, and SASS teacher control number.

After teachers are sorted using the above variables, they are selected within each stratum using a systematic probability proportional to size (PPS) sampling procedure. Any teacher with a measure of size greater than the sampling interval is included in the sample with certainty (i.e., automatically included). Since TFS selection probabilities are not conditioned on anything, the selected sample sizes equal the allocated sample size.

The 2008–09 TFS sample consisted of about 5,500 teachers out of the 57,000 public and private school teachers who participated in the 2007–08 SASS. (See chapter 4 for information on the SASS sample design.)

Data Collection and Processing

The 2008–09 TFS data collection was an online collection, followed by e-mail and telephone reminders, a hard-copy mailing, and telephone follow-up. The U.S. Census Bureau is the data collection agent.

Reference dates. Most data items refer to teacher status at the time of questionnaire completion. Some items refer to the past school year, the past 12 months, or the next school year.

Data collection. In the fall of the year of the survey administration, the Census Bureau mails a Teacher Status Form to each school that had at least one teacher who participated in the previous year's SASS. On this form, the school principal (or other knowledgeable staff member) is asked to report the current occupational status of each teacher listed by indicating whether that teacher (1) is still at the school in a teaching or nonteaching capacity; or (2) has left the school to teach elsewhere or to enter a nonteaching occupation. If school staff indicates that a sample teacher has moved, and the teacher did not provide contact information on his or her SASS questionnaire, the Census Bureau tries to obtain the correct home address from the U.S. Postal Service.

For the 2008–09 TFS, the link to the user IDs and passwords for access to the online questionnaire were mailed to selected SASS teachers in early February 2009. The letters were mailed to home addresses, where available; otherwise, they were mailed to the sample teacher's school as listed in the previous SASS administration.

In March 2009, Census interviewers began calling sampled teachers who had not yet completed the survey. If the interviewers were unable to contact a sampled teacher through a contact person or through directory assistance, they called the sampled teacher's school to obtain information about his or her current address or employer. Interviewers used the same online instrument to collect the data as was used by the sampled teachers to complete the survey. Teachers who had not completed the online instrument as of April 2009 were sent a hard-copy version of the questionnaire.

Editing. Surveys undergo several stages of editing. TFS data that were provided on hard-copy versions of questionnaires are converted from paper to electronic format using manual data keying. All keyed entries are 100 percent verified by the keying staff, meaning that each field is keyed twice and the results are compared automatically for discrepancies and, subsequently, verified. All survey data are then reformatted into SAS datasets in order to begin the extensive preliminary data review process. During this stage, analysts split the TFS data into two files: a former teacher file (for leavers) and a current teacher file (for stayers and movers).

The next step is to make a preliminary determination of each case's interview status recode (ISR) value; that is, whether it is an interview, a noninterview, or out-of-scope for the survey. Records classified as interviews are submitted to a series of computer edits: range checks, consistency edits, and blanking edits. Next, the records undergo a final edit to determine whether the case is eligible to be included in the survey and, if so, whether sufficient data have been collected for the case to be classified as a completed interview. A final ISR value is then assigned to each case as a result of this edit.

Estimation Methods

Estimates from TFS sample data are produced using weighting and imputation procedures.

Weighting. The general purpose of weighting is to scale up the sample estimates to represent the target survey population. In TFS, the steps for weighting types of respondents are similar to those used for SASS. For TFS, a base weight (the inverse of the sampled teacher's probability of selection) is used as the starting point. Then, a weighting adjustment is applied that reflects the impact of the SASS teacher weighting procedure. Next, a nonresponse adjustment factor is calculated and applied using information known about the respondents from the sampling frame data. Finally, a ratio adjustment factor is calculated and applied to the sample to adjust the sample totals to frame totals in order to reduce sampling variability. The product of these factors is the final weight for TFS.

Imputation. In all administrations of TFS, all items missing values are imputed for records classified as interviews. In order to fill these items with data, questionnaires are put through three independent stages of imputation. The first stage involves using items from the same TFS questionnaire or items from the corresponding SASS school or teacher questionnaire to impute the missing data. In the second stage, any remaining unanswered items are imputed using "hot-deck" imputation (in which donor records are established and used to impute data). In the third and final stage, any remaining unanswered items are imputed clerically by Census Bureau analysts. The third stage is necessary when there is no available donor or the value imputed by computer is inconsistent with values in other items.

Future Plans

SASS is now conducted on a 4-year cycle, with the next collection planned for the 2011–12 school year. TFS is also conducted on a 4-year cycle (in the school year following the SASS administration). The

next TFS administration is scheduled for the 2012–13 school year.

5. DATA QUALITY AND COMPARABILITY

Sampling Error

Because the TFS sample is a subsample of the SASS teacher sample, the SASS teacher replicate weights are used to derive the TFS replicate weights. (See the discussion of sampling error and variance estimation for SASS in chapter 4.) The base weight for each TFS teacher is multiplied by each of the SASS replicate weights divided by the SASS teacher full-sample base weight for that teacher. To calculate the 88 replicate weights, which should be used for variance calculations, these TFS replicate basic weights are processed through the remainder of the TFS weighting system.

Nonsampling Error

Coverage error. There is a potential for bias to be introduced into TFS because the TFS frame only includes teachers who responded to SASS.

Nonresponse error.

Unit nonresponse. The total weighted unit response rate in the 2008–09 TFS was 88 percent. The weighted response rate for former teachers (who completed the Former Teacher Questionnaire) was slightly lower than the weighted response rate for current teachers (who completed the Current Teacher Questionnaire) (85 vs. 88 percent, respectively).

The overall response rate represents the response rate to the survey, taking into consideration each stage of data collection. For a teacher to be eligible for TFS, it was necessary to have received the Teacher Listing Form from the school during the previous year's SASS data collection, which provided a sampling frame for teachers at that school, and for the teacher to have responded to the SASS teacher questionnaire. The overall response rate (shown in Table 4) is calculated as follows: SASS Teacher Listing Form response rate x SASS teacher questionnaire response rate x TFS questionnaire response rate.

Item nonresponse. Item response rates indicate the percentage of respondents who answered a given survey question or item. The weighted TFS item response rates are produced by dividing the number of sampled teachers who responded to an item by the number of sampled teachers who were eligible to

Table 4. Base-weighted response rates for SASS teacher data files and TFS data files, by sector: School years 2007-08 and 2008-09

Sector	Base-weighted 2007-08 SASS Teacher Listing Form response rate	Base-weighted 2007-08 SASS teacher data file response rate	Base-weighted 2008-09 TFS response rate		Overall 2008-09 TFS response rate	
			Current teachers	Former teachers	Current teachers	Former teachers
Total	85.9	83.3	88.3	84.7	63.2	60.6
Public ¹	86.2	84.0	88.4	84.8	64.0	61.4
Private	85.1	77.5	87.1	84.4	57.4	55.7

¹The public sector includes teachers from traditional public and public charter schools.

NOTE: Base-weighted response rates use the inverse of the probability of selection and the sampling adjustment factor.

SOURCE: Cox, S., Parmer, T., Tourkin, S., Warner, T., and Lyter, D.M. (2007). *Documentation for the 2004–05 Teacher Follow-up Survey* (NCES 2007-349). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

answer that item, and then adjusting those rates by the final weight. In the 2008–09 TFS, the weighted item response rates for the Former Teacher Questionnaire ranged from 75 to 100 percent. The weighted item response rates for the Current Teacher Questionnaire ranged from 74 to 100 percent. The Former Teacher Questionnaire had six items that had a weighted response rate of less than 85 percent. The Current Teacher Questionnaire had four items that had a weighted response rate of less than 85 percent.

Measurement error. Reinterviews were conducted for the purpose of measuring response variance in the 1994–95 TFS. The reinterviews were conducted through two reinterview questionnaires—one for mail cases and another for telephone cases. Each questionnaire contained a subset of questions from the original questionnaire. Seventy-eight percent of the questions evaluated displayed high response variance; only 5 percent displayed low response variance. (All but one of the 54 questions on teaching methods had moderate or high response variance.) This reinterview study again confirmed that “mark all that apply” questions tend to be problematic. See *Response Variance in the 1994–95 Teacher Follow-up Survey* (Bushery et al. 1998).

Data Comparability

Care must be taken in estimating change over time in a TFS data element, because some of the measured change may not be attributable to a change in the educational system, but due to changes in the sampling frame, questionnaire item wording, or other changes. For example, the definitions of the locale codes based on the U.S. Census were revised in 2000 and again in 2003. Changes in how schools’ locales are categorized over time may account for at least

some changes that are noted from previous administrations. This impacts the urbanicity variables included in the data files, which are based on the 2000 Census definitions for locale codes.

For further information on the comparability of data elements, see Appendix M in *Documentation for the 2008–09 Teacher Follow-up Survey* (forthcoming). Appendix M contains crosswalks that compare items in the 2008–09 TFS with items in the 2000–01 TFS and the 2007–08 SASS Teacher Questionnaire.

6. CONTACT INFORMATION

For content information on the TFS project, contact:

Freddie Cross
Phone: (202) 502-7489
E-mail: freddie.cross@ed.gov

Kerry Gruber
Phone: (202) 502-7349
E-mail: kerry.gruber@ed.gov

SASS e-mail: sassdata@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Cox, S., Parmer, T., Tourkin, S., Warner, T., and Lyter, D.M. (2007). *Documentation for the 2004–05 Teacher Follow-up Survey* (NCES 2007-349). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Cox, S., Parmer, T., Chambers, L., Tourkin, S., and Lyter, D. (forthcoming). *Documentation for the 2008–09 Teacher Follow-up Survey: Public-Use Version*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Tourkin, S.C., Warner, T., Parmer, R., Cole, C., Jackson, B., Zukerberg, A., Cox, S., and Soderborg, A. (2007). *Documentation for the 2003–04 Schools and Staffing Survey* (NCES 2007-337). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Whitener, S.D. (1994). *1991–92 SASS Teacher Follow-up Survey Data File User's Manual Public-Use Version* (NCES 94-331). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Whitener, S.D. (1994). *1991–92 SASS Teacher Follow-up Survey Data File User's Manual Restricted-Use Version* (NCES 94-478). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Whitener, S.D., Gruber, K.J., Rohr, C., and Fondelier, S. (1998). *1994–95 Teacher Follow-up Survey, Data File User's Manual Public-Use Version* (NCES 98-232). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Whitener, S.D., Gruber, K.J., Rohr, C.L., and Fondelier, S.E. (1999). *1994–95 Teacher Follow-up Survey, Data File User's Manual Restricted-Use Codebook* (NCES 1999-14). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Bushery, J.M., Schreiner, I.D., and Newman-Smith, A. (1998). *Response Variance in the 1994–95 Teacher Follow-up Survey* (NCES 98-13). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Jenkins, C.R., and Wetzel, A. (1995). *The Results of the 1991–92 Teacher Follow-up Survey (TFS) Re-interview and Extensive Reconciliation* (NCES 95-10). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Mullens, J.E. (1995). *Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey* (NCES 95-15). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Salvucci, S., Walter, E., Conley, V., Fink, S., and Newman-Smith, A. (1997). *Measurement Error Studies at the National Center for Education Statistics* (NCES 97-464). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Chapter 6: National Longitudinal Study of the High School Class of 1972 (NLS:72)

1. OVERVIEW

In response to the need for policy-relevant, time-series data on nationally representative samples of elementary and secondary students, NCES instituted the National Longitudinal Studies (NLS) Program, a continuing long-term project. The general aim of this program is to study the educational, vocational, and personal development of students at various grade levels and the personal, familial, social, institutional, and cultural factors that may affect that development. The National Longitudinal Study of the High School Class of 1972 (NLS:72) was the first in the series. The first four studies—NLS:72, the High School and Beyond Longitudinal Study (HS&B) (see chapter 7), the National Education Longitudinal Study of 1988 (NELS:88) (see chapter 8), and the Education Longitudinal Study of 2002 (ELS:2002) (see chapter 9)—cover the educational experience of youth from the 1970s into the 21st century.

NLS:72 collected comprehensive base-year data from a nationally representative sample of high school seniors in spring 1972, prior to high school graduation. Additional information about students and schools was obtained from school administrators and counselors. Over the course of the project—extending from the base-year survey in 1972 to the fifth follow-up survey in 1986—data were collected on nearly 23,000 students. A number of supplemental data collection efforts were also undertaken, including a Postsecondary Education Transcript Study (PETS) in 1984–85 and a Teaching Supplement in 1986.

Purpose

To provide information on the transition of young adults from high school through postsecondary education and into the workplace.

Components

NLS:72 collected data from students (high school seniors in 1972), school administrators, and school counselors. Data were primarily collected in a base-year and five follow-up surveys. The project also included periodic supplements completed by 1972 high school seniors and a collection of postsecondary transcripts from the colleges and universities attended by the students.

Base-year survey. The base-year survey was conducted in spring 1972 and comprised the following:

Student Questionnaire. Students reported information about their personal and family background (age, sex, race, physical handicaps, socioeconomic status [SES] of family and community); education and work experiences (school characteristics and performance; work status, performance, and satisfaction); future plans (work, education, and/or military); and aspirations, attitudes, and opinions. Students also completed a *Test Battery*—six timed aptitude tests that measured verbal and nonverbal abilities. These tests covered vocabulary, picture number, reading, letter groups,

LONGITUDINAL SAMPLE SURVEY OF THE HIGH SCHOOL SENIOR CLASS OF 1972. BASE-YEAR SURVEY AND FIVE FOLLOW-UPS, ENDING IN 1986

NLS:72 collected data from:

- Students
- School administrators
- School counselors
- Postsecondary transcripts

mathematics, and mosaic comparisons. (See “Test Battery” in Section 3. Key Concepts.)

Student Record Information Form (SRIF). School administrators completed this form for each student sample member. The SRIF collected data on each student’s high school curriculum, credit hours in major courses, and grade point average (and, if applicable, the student’s position in ability groupings, remedial-instruction record, involvement in certain federally supported programs, and scores on standardized tests).

School Questionnaire. School administrators provided data on program and student enrollment information, such as grades covered, enrollment by grade, curricula offered, attendance records, racial/ethnic composition of school, dropout rates by sex, number of handicapped and disadvantaged students, and percentage of recent graduates in college.

Counselor Questionnaire. One or two counselors in each school provided data on their sex, race, and age; college courses in counseling and practice background; total years of counseling and years at present school; prior counseling experience with Black, Hispanic, and other race/ethnicity groups; sources of support for postsecondary education recommended to/used by students; job placement methods used; number of students assigned for counseling and number counseled per week; time spent in counseling per week; time spent with students about various problems, choices, and guidance; and time spent in various other activities (e.g., conferences with parents and teachers).

Follow-up surveys. In 1973, 1974, 1976, 1979, and 1986, NCES conducted follow-up surveys of students in the 1972 base-year sample and of students in an augmented sample selected for the first follow-up. These surveys collected information from the 1972 high school seniors on marital status; children; community characteristics; education, military service, and/or work plans; educational attainment (schools attended, grades received, credits earned, financial assistance); work history; attitudes and opinions relating to self-esteem, goals, job satisfaction, and satisfaction with school experiences; and participation in community affairs or political activities. School Questionnaires and retrospective high school data were collected during the first follow-up for sample schools and students who had not participated in the base-year survey.

Concurrently with the second follow-up, an *Activity State Questionnaire* was administered to sample

members who had not provided activity information in the base-year or first follow-up surveys. Data were collected on pursuits in which the sample member was active in October of 1972 and 1973, including education, work, military service, and being a housewife, among others. Background information about the sample member’s high school program and about parents’ education and occupation was also requested.

During the fourth follow-up survey, a subsample of respondents was retested on a subset of the base-year Test Battery. In addition, a *Supplemental Questionnaire* was administered to respondents who had not reported certain information in previous surveys. The information asked for retrospectively covered the sample member’s school and employment status from October 1972 through October 1976 and his or her license or diploma status as of October 1976. The questionnaires were tailored to the sample member’s pattern of missing responses and consisted of two to four of the 11 possible sections.

The fifth follow-up survey offered the opportunity to gather information on the experiences and attitudes of sample members for whom an extensive history already existed. It differed from the previous follow-ups in that it was only sent to a subsample of the original respondents and targeted certain subgroups in the population. About 10 pages of new questions on marital history, divorce, child support, and economic relationships in families were included. The fifth follow-up also included a sequence of questions aimed at understanding the kinds of individuals who apply for and enroll in graduate management programs, as well as several questions about attitudes toward the teaching profession.

A *Teaching Supplement*, which was administered concurrently with the fifth follow-up, was a separate questionnaire that was sent to fifth follow-up respondents who indicated on the main survey form that they had teaching experience or training. The supplement focused on the qualifications, experiences, and attitudes of current and former elementary and secondary school teachers and on the qualifications of persons who had completed a degree in education or who had received certification, but had not actually taught. The supplement included items that asked about reasons for entering the teaching career, degrees and certification, actual teaching experience, allocation of time while working, pay scale, satisfaction with teaching, characteristics of the school in which the respondent taught, and professional activities. Former teachers were asked about their reasons for leaving the teaching profession and the career (if any) they pursued

afterward. Current teachers were asked about their future career plans, including how long they expected to remain in teaching. The supplement included six critical items: type of certification, certification subject(s), first year of teaching, beginning salary in the district where the respondent was currently teaching, years of experience, and grade level taught.

Postsecondary Education Transcript Study (PETS).

To obtain data on coursework and credits for analysis of occupational and career outcomes, NCES requested official transcripts from all academic and vocational schools attended by the 1972 seniors since leaving high school. This study, conducted during 1984–85, collected transcripts from all postsecondary institutions reported by sample members in the first through fourth follow-up surveys. The information gathered from the transcripts included terms of attendance, fields of study, specific courses taken, and grades and credits earned. As the study covered a 12-year period, dates of attendance and term dates were recorded from each transcript received, allowing analysis over the whole period or any defined part.

Periodicity

The base-year survey was conducted in the spring of 1972, with five follow-ups in 1973, 1974, 1976, 1979, and 1986. Supplemental data collections were administered during all but the third follow-up. Postsecondary transcripts were collected in 1984–85.

2. USES OF DATA

NLS:72 is the oldest of the longitudinal studies sponsored by NCES. It is probably the richest archive ever assembled on a single generation of Americans. Young people's success in making the transition from high school or college to the workforce varies enormously for reasons only partially understood. NLS:72 data can provide information about the quality, equity, and diversity of educational opportunity and the effect of these factors on cognitive growth, individual development, and educational outcomes. It can also provide information about changes in educational and career outcomes and other transitions over time.

The Teaching Supplement data can be used to investigate policy issues related to teacher quality and retention. These data can be linked to data from prior waves of the Student Questionnaire for analysis of antecedent conditions and events that may have influenced respondents' career decisions. The data can also be merged with results from the fifth follow-up

questionnaire, which included special questions related to teaching.

The history of the members of the class of 1972, from their high school years through their early 30s, is widely considered as the baseline against which the progress and achievements of subsequent cohorts are to be measured. Researchers have drawn on this archive since its inception. To date, the principal comparisons have been with the other three longitudinal studies: HS&B, NELS:88, and ELS:2002. Together, these four studies provide a particularly rich resource for examining the changes that have occurred in American education during the past 30 years. Data from these studies can be used to examine how student academic coursework, achievement, values, and aspirations have changed, or remained constant, throughout this period.

The NLS studies offer a number of possible time points for comparison. Cohorts can be compared on an *intergenerational or cross-cohort time-lag basis*. Both cross-sectional and longitudinal time-lag comparisons are possible. For example, cross-sectionally, NLS:72 seniors in 1972 can be compared to HS&B base-year seniors in 1980, NELS:88 second follow-up seniors in 1992, and ELS:2002 first follow-up seniors in 2004. Longitudinally, changes measured between the senior year and 2 years after graduation can be compared across studies. *Fixed time comparisons* are also possible; groups within each study can be compared to each other at different ages, but at the same point in time. Thus, NLS:72 seniors, HS&B seniors, and HS&B sophomores can all be compared in 1986—some 14, 6, and 4 years after each respective cohort completed high school. Finally, *longitudinal comparative analyses* of the cohorts can be performed by modeling the history of the age/grade cohorts. The possible comparison points and the considerations of content and design that may affect the comparability of data across the cohorts are discussed in *Trends Among High School Seniors, 1972–1992* (Green, Dugoni, and Ingels 1995) and *United States High School Sophomores: A Twenty-Two Year Comparison, 1980–2002* (Cahalan et al. 2006).

3. KEY CONCEPTS

A few key terms relating to NLS:72 are defined below.

Test Battery. Six cognitive tests were administered during the base year: (1) vocabulary (15 items, 5 minutes), a brief test using a synonym format; (2) picture number (30 items, 10 minutes), a test of associative memory consisting of a series of drawings of familiar objects, each paired with a number; (3) reading (20 items, 15

minutes), a test of comprehension of short passages; (4) letter groups (25 items, 15 minutes), a test of inductive reasoning that required the student to draw general concepts from sets of data or to form and try out hypotheses in a nonverbal context; (5) mathematics (25 items, 15 minutes), a quantitative comparison in which the student indicated which of two quantities was greater (or asserted their equality or the lack of sufficient data to determine which quantity was greater); and (6) mosaic comparisons (116 items, 9 minutes), a test measuring perceptual speed and accuracy through the use of items that required detection of small differences between pairs of otherwise identical mosaic, or tile-like, patterns.

Socioeconomic status (SES). A composite scale developed as a sum of standardized scales of father's education, mother's education, 1972 family income, father's occupation, and household items. The latter two underlying scales were computed from base-year Student Questionnaire responses. The other three underlying scales were derived from base-year responses as augmented by first follow-up responses and responses to a second follow-up resurvey in order to obtain this and other information from sample members who had failed to provide it previously. Each index component was first subjected to factor analysis that revealed a common factor with approximately equal weights for each component. Each of the components was then standardized, and an equally weighted combination of the five standard scores yielded the SES composite score. The data file contains both the raw score and a categorized SES score (SES Index).

4. SURVEY DESIGN

Target Population

The population of students who, in spring 1972, were 12th graders (high school seniors) in public and private schools located in the 50 states and the District of Columbia. Excluded were students in schools for the physically or mentally handicapped, students in schools for legally confined students, early (mid-year) graduates, dropouts, and individuals attending adult education classes.

Sample Design

Base-year survey. The NLS:72 sample was designed to be representative of the approximately 3 million high school seniors enrolled in more than 17,000 schools in the United States in spring 1972. The base-year sample design was a stratified, two-stage probability sample of students from all public and private schools in the 50 states and the District of Columbia that enrolled 12th graders in the 1971–72 school year. Excluded were

schools for the physically or mentally handicapped and schools for legally confined students. A sample of schools was selected in the first stage. In the second stage, a random sample of 18 high school seniors was selected within each participating school.

The base-year first-stage sampling frame was constructed from computerized school files maintained by the U.S. Department of Education and the National Catholic Educational Association. The original sampling frame called for 1,200 schools; that is, 600 strata with two schools per stratum. The strata were defined based upon the following variables: type of control (public or private), geographic region, grade 12 enrollment size, geographic proximity to institutions of higher education, proportion of Black, Hispanic, and other race/ethnicity student enrollment (for public schools only), income level of the community, and degree of urbanization. Schools were selected with equal probability for all but the smallest size stratum (schools with enrollment under 300). In that stratum, schools were selected with probability proportional to enrollment. All selections were without replacement. To produce sufficient sizes for intensive study of disadvantaged students, schools in low-income areas and schools with high proportions of Black, Hispanic, and other race/ethnicity student enrollment were sampled at twice the rate used for the remaining schools. Within each stratum, four schools were selected, and then two of the four were randomly designated as the primary selections. The other two schools were retained as backup or substitute selections (for use only if one or both of the primary schools did not cooperate).

The second stage of the base-year sampling procedure consisted of first drawing a simple random sample of 18 students per school (or all students, if fewer than 18 were available) and then selecting 5 additional students (if available) as possible substitutes for nonparticipants. In both cases, the students within a school were sampled with equal probability and without replacement. Dropouts, early (mid-year) graduates, and individuals attending adult education classes were excluded from the sample. The oversampling of schools in low-income areas and schools with relatively high Black, Hispanic, and other race/ethnicity student enrollment led to oversampling of low-income and Black, Hispanic, and other race/ethnicity students.

Sample redefinitions and augmentations. At the close of the base-year survey, 1,040 schools (950 primary and 100 backup) of a targeted 1,200 schools and 26 “extra” backup schools had participated (school participation being defined as students from that school contributing SRIFs, Test Batteries, or Student Questionnaires). A backup school was termed “extra” if, ultimately, both

primary sample schools from that stratum also participated. An additional 21 primary schools indicated that they had no 1972 seniors. At this point, there remained several strata with no participating schools and many more with only one school. To reduce the effects of the large base-year school nonresponse, a resurvey activity was implemented in the summer of 1973 (prior to the first follow-up survey). An attempt was made to elicit cooperation from the 231 nonparticipating base-year primary sample schools and to obtain backup schools to fill empty or partially filled strata. The resurvey was successful in 205 of the 231 primary sample schools. Students from 36 backup schools were also included in order to obtain at least two participating schools in the first follow-up survey from each of the 600 original strata. Students from the 26 “extra” backup schools from the base-year survey were not surveyed during the first follow-up; however, students from 18 of these schools were included in the second and subsequent follow-up surveys to avoid elimination of cases with complete base-year data.

To compensate for base-year school undercoverage, samples of former 1972 high school seniors were selected for inclusion in the first and subsequent follow-ups from 16 sample augmentation schools (8 new strata); these schools were selected from those identified in 200 sample school districts canvassed to identify public schools not included in the original sampling frame. As before, 18 students per school were selected (as feasible) by simple random sample.

The number of students in the final sample from each sample school was taken as the number of students who were offered a chance to be in the sample and were eligible for the study. This included both respondents and nonrespondents, but excluded ineligible students, such as dropouts, early (mid-year) graduates, and those attending adult education classes. The final NLS:72 sample included 23,450 former 1972 high school seniors and 1,340 sample schools—1,150 participating primary schools, 21 primary schools with no 1972 seniors, 131 backup sample schools, 18 “extra” schools in which base-year student data had been completed, and 16 augmentation schools.

A subsample of 1,020 of the 14,630 eligible fourth follow-up sample members (those who had completed both a Student Questionnaire and a Test Battery in the base-year survey) was targeted for retests on a subset of the base-year Test Battery. Because a self-weighting subsample would have yielded an inadequate number of Black subsample members, a design option that oversampled Blacks was adopted. In addition to the stratification by race, the sample was controlled within strata on three factors believed to be highly correlated

with retest ability scores: base-year ability, SES, and postsecondary educational achievement. The control was achieved by applying an implicit stratification procedure. Test results were obtained from 692 of those in the subsample. Additional retest data were requested for all fourth follow-up sample members who had participated in the base-year testing and who were scheduled for a personal interview. This resulted in additional test data for 1,960 individuals (50.3 percent of those defined as retest-eligible).

Fifth follow-up survey. The fifth follow-up sample was an unequal probability subsample of the 22,650 students who had participated in at least one of the five previous waves of NLS:72. The fifth follow-up retained the essential features of the initial stratified multistage design but differed from the base-year design in that the secondary sampling unit selection probabilities were unequal, whereas they were equal in the base-year design. This inequality of selection probabilities allowed oversampling of policy-relevant groups and enabled favorable cost-efficiency tradeoffs.

In general, the retention probabilities for students were inversely proportional to the initial sample selection probabilities. The exceptions were for (1) sample members with special policy relevance, who were retained with certainty or at a higher rate than other sample members; (2) persons with very small initial selection probabilities, who were retained with certainty; and (3) nonparticipants in the fourth follow-up, who were retained at a lower rate than other sample members because they were expected to be more expensive to locate and because they would be less useful for longitudinal analysis.

The subgroups of the original sample retained with certainty were (1) Hispanics who participated in the fourth follow-up survey; (2) teachers and “potential teachers” who participated in the fourth follow-up survey (a “potential teacher” was defined as a person who majored in education in college or was certified to teach or whose background was in the sciences); (3) persons with a 4-year or 5-year college degree or a more advanced degree; and (4) persons who were divorced, widowed, or separated from their spouses, or never-married parents. These groups overlapped and did not comprise distinct strata in the usual sense.

Teaching Supplement. The fifth follow-up sample included all sample members known to be teachers or potential teachers as of the fourth follow-up in 1979. To identify those sample members who had become teachers between the fourth and fifth follow-ups, a direct question was included in the fifth follow-up main questionnaire. Sample members were selected for the

Teaching Supplement sample if they indicated that they were (1) currently an elementary or secondary teacher; (2) formerly an elementary or secondary teacher; or (3) trained as an elementary or secondary teacher but never went into teaching. Of the 12,840 fifth follow-up respondents, 1,520 were eligible for the Teaching Supplement.

Postsecondary Education Transcript Study (PETS). In the first through fourth follow-up surveys, approximately 14,700 members of the NLS:72 cohort reported enrollment at one or more postsecondary institutions. An attempt was made to obtain a transcript from each school named by a respondent. Thus, no probabilistic sampling was done to define the PETS sample.

Data Collection and Processing

The base-year survey was administered through group administration. For the first four follow-up surveys, field operations began in the summer or fall of the survey year and continued through the spring of the following year; for example, the third follow-up survey data collection began in October 1976 and continued through June 1977. For the fifth follow-up survey, the data collection began in March 1986 and ended in mid-September 1986. The Educational Testing Service (ETS) administered the base-year survey; the Research Triangle Institute (RTI) carried out the first through fourth follow-up surveys; and the National Opinion Research Center (NORC) conducted the fifth follow-up survey.

Reference dates. Sample members in each of the first four follow-up surveys were asked about their family (marital status, spouse's status, number of children), location, and what they were doing with regard to work, education, and/or training during the first week of October of the survey year; fifth follow-up participants were asked the same questions for the first week of February 1986. Family income was requested for the preceding 2 years, and political and volunteer activities were requested for the past 24 months. Participants in each follow-up survey were also asked for summaries of educational and work experiences and activities for the intervening year(s) since the last survey. For the first four follow-up surveys, this information was requested as of the month of October in the intervening year(s) or sometimes overall for each year preceding the survey; fifth follow-up survey participants were asked detailed questions for up to four jobs and for attendance at up to two educational institutions since October 1979.

Data collection. Data collection instruments and procedures for the base-year survey were designed during the 1970–71 school year and were tested on a small sample of high school seniors in spring 1971. One

year later, the full-scale NLS:72 study was initiated. Through an in-school group administration in the base year, each student was asked to complete a Test Battery (measuring both verbal and nonverbal aptitude) and applicable portions of a Student Questionnaire containing 104 questions distributed over 11 major sections. Students were given the option of completing the Student Questionnaire in school or taking it home and answering the questions with the assistance of their parents. In addition, school administrators at each participating school were asked to complete a School Questionnaire and an SRIF for each student in the sample. One or two counselors from each school in the sample were asked to complete a Counselor Questionnaire.

Follow-up surveys. In fall 1973, 1974, 1976, and 1979 and spring 1986, sample members (or a subsample) were again contacted. After extensive tracing to update the name and address files, follow-up questionnaires were mailed to the last known addresses of sample members whose addresses appeared sufficient and correct and who had not been removed from active status by prior refusal, reported death, or other reason. Respondents to the third through fifth follow-ups were offered small monetary incentives for completing the questionnaires. The mailouts were followed by a planned sequence of reminder postcards; additional questionnaire mailings; reminder mailgrams (for the first four follow-ups) and telephone calls; personal interviews; and, for the third to fifth follow-ups only, telephone interviews of nonrespondents. During personal interviews, the entire questionnaire was administered. During the telephone interviews conducted in the last three follow-ups, only critical items that were suitable for telephone administration were administered. In order to make survey procedures comparable, respondents were asked to keep a copy of the questionnaire in front of them for both telephone and in-person interviews.

In all follow-ups, returned questionnaire cases missing critical items were flagged during data entry, and data were retrieved by specially trained telephone interviewers. Although most questions were of the forced-choice type, coding was required for the open-ended questions on occupation, industry, postsecondary school, field of study, state where marriage and divorce occurred, and relationship. Occupational and industry codes were obtained from the U.S. Census Bureau's *Classified Index of Industries and Occupations, 1970* and *Alphabetical Index of Industries and Occupations, 1970*. These sources were used in all follow-ups. Coding of the names of postsecondary schools attended by respondents was accomplished using codes from NCES's *Education Directory, Colleges and Universities*. Field of study information was coded using classification of

instructional program (CIP) codes from NCES's *Classification of Instructional Programs*. In the fifth follow-up, for the first time, all codes were loaded into a computer program for quicker access. Coders entered a given response, and the program displayed the corresponding numerical code.

Prior to the fifth follow-up, all data were entered via direct access terminals. The fifth follow-up survey marked the first time that NLS:72 data were entered with a combination of keyed entry and optical scanning procedures. Using a computer-assisted data entry (CADE) system, operators were able to combine data entry with traditional editing procedures. All critical items and filter items (plus error-prone data like dollar amounts and numbers in general) were processed by CADE. The remaining data were optically scanned.

Teaching Supplement. Data collection procedures used for the Teaching Supplement, administered concurrently with the fifth follow-up, were similar to those used for the follow-up surveys.

Postsecondary Education Transcript Study (PETS). Packets of transcript survey materials were mailed to the postsecondary schools in July 1984, with a supplemental mailing in November 1984. Altogether, 24,430 transcripts were initially requested from 3,980 institutions for 14,760 NLS:72 sample members. Telephone follow-up of nonresponding schools began in September 1984, when transcripts had been received from about two-thirds of the schools.

After investigating several alternatives, NORC adapted its CADE system for processing postsecondary transcripts. A single member of the specially trained data preparation staff analyzed the transcript document to determine its general organization and special characteristics; abstracted standard information from the document into a common format; assigned standard numerical codes to such transcript data elements as major and minor fields of study, degrees earned, types of academic term, titles of courses taken, and grades and credits; and entered all pertinent information into a computer file. Combining these steps ensured that transcripts would be handled as internally consistent, integrated records of an individual's educational activity. Moreover, since all transcript processing occurred at a single station, the use of CADE reduced the number of steps at which records might be lost or misrouted or other errors introduced into the database.

Editing. For the base-year through fourth follow-up surveys, an extensive manual or machine edit of all NLS:72 data was conducted in preparing the release file for public use. Editing involved rigorous consistency

checking of all routing patterns within an instrument (not just skip patterns containing "key" or critical items), as well as range checks for all items and the assignment of error or missing data codes as necessary. Checks of the hard-copy sources were required in some cases for error resolution.

Unlike the earlier surveys, all editing for the fifth follow-up was carried out as part of CADE. The machine-editing steps used in the prior follow-ups were implemented for scanned items. Since most of the filter questions in the fifth follow-up were CADE-designated items, there were few filter-dependent inconsistencies to be handled in machine editing. Validation procedures for the fifth follow-up centered on verification of data quality through item checks and verification of the method of administration for 10 percent of each telephone or personal interviewer's work. Field managers telephoned the respondent to check several items of fact and to confirm that the interviewer had conducted a personal or telephone interview or had picked up a questionnaire. No cases failed validation.

Postsecondary Education Transcript Study (PETS). The CADE system enforced predetermined range and value limitations on each field. It performed three types of error screenings: (1) a check-digit system, which disallowed entry of incorrect identification data (school codes from the Federal Interagency Committee on Education (FICE), student identification numbers, and combinations of schools and students); (2) each data field was programmed to disallow entry of illogical or otherwise incorrect data; and (3) each CIP code selected to classify a field of study or a course was confirmed by automatically displaying the CIP program name for the code next to the name (from the original CADE transcript) that the coder had entered. A sample of CADE transcripts was selected and printed from every completed data disk for supervisory review.

Estimation Methods

Data were weighted in NLS:72 to adjust for sampling and nonresponse. Various composite variables have also been computed to assist in data analyses.

Weighting. The weighting procedures used for the various NLS:72 survey data are described below.

Student files. NLS:72 student weights are based upon the inverse of the probabilities of selection through all stages of the sampling process and upon nonresponse adjustment factors computed within weighting classes. Unadjusted raw weights—the inverses of sample inclusion probabilities—were calculated for all students sampled in each survey year. These weights are a function of the school selection probabilities and the

student selection probabilities within a school. The raw weight for a case equals the raw weight for the base-year sample divided by the conditional probability of selection into that follow-up survey, given that the case was selected into the base-year sample.

Because of the various sample redefinitions and augmentations and nonresponse to the various student instruments, several sets of adjusted weights were computed for each NLS:72 survey wave. Each weight is appropriate for a particular respondent group. The general adjustment procedure used was a weighting class approach, which distributes the weights of nonrespondents to respondents who are in the same weighting class. The adjustment involves partitioning the entire student sample (respondents and nonrespondents) into weighting classes (homogeneous groups with respect to survey classification variables) and performing the adjustments within weighting class. Adjusted weights for nonrespondents are set to 0, and their adjusted weights are distributed to respondents proportionally to the respondents' unadjusted weights. Differential response rates for students in different weighting classes are reflected in the adjustment, and the weight total within each weighting class (and thus for the sample as a whole) is maintained.

The weighting class cells were defined by cross-classifying cases by several variables. For the first through fourth follow-up surveys, the weighting class cells were sex, race, high school program, high school grade point average, and parents' education. For the fifth follow-up survey, the weighting class cells were similar except that postsecondary education attendance was substituted for parents' education. In some instances, cells were combined by pooling across certain weighting class cells.

The adjusted weights in the third and fourth follow-ups are applicable only to key items in these surveys (or specified combinations of these items with items from other instruments). The restriction is related to a change in data collection procedures. One or two item nonresponse adjustment factors were calculated for each of these surveys for the nonkey items that were not asked. The appropriate adjusted weight for each survey should be multiplied by its nonresponse adjustment factor to provide a new weight that is appropriate to items in that survey that are not key (or combinations of such nonkey items with items from other instruments).

Refer to the NLS:72 *Fifth Follow-Up (1986) Final Technical Report* (Sebring et al. 1987) for complete weighting procedures and a specification of available weights and appropriate variables to which the weights apply.

Teaching Supplement file. One set of weights was specifically developed to compensate for the unequal probabilities of retention in the Teaching Supplement sample and to adjust for unit nonresponse. Theoretically, the weights project to the population of high school seniors of 1972 who have taught elementary or secondary school or who were trained to teach but never went into teaching. The weighting procedures were similar to those used in the follow-up surveys and consisted of two basic steps. The first step was the calculation of a preliminary weight based on the inverse of the cumulative probabilities of selection for the Teaching Supplement. The preliminary weight for the Teaching Supplement is the fifth follow-up adjusted weight. The second step carried out the adjustment of this preliminary weight to compensate for unit nonresponse. Respondents were cross-classified into weighting cells by race, high school grades, and status as a teacher (current or former teacher, or never taught).

School file. During the sequential determination of final school sample membership (including augmentations), several school sample weights were computed. The principal purpose of the various school weights was to serve as a basis for the subsequent computation of student weights applicable to one or more of the student instruments. Only two of the eight weights computed are of direct use in analyzing school file or other school-level data. The school file sample weight is appropriate for analyzing school-level data that potentially could be supplied by all schools, including the School Questionnaire data.

The adjusted counselor weight should be used only in analyzing the responses to the Counselor Questionnaire; however, care must be exercised when analyzing these data. This questionnaire was only administered at base-year responding schools, and data were collected from either one or two counselors at each school.

Postsecondary Education Transcript Study (PETS) file. Because the PETS did not introduce any additional subsampling into the NLS:72 sample design, it was not necessary to calculate a new raw weight for this study. Instead, the raw weight for the base-year survey was used to create three adjusted weights specifically for the analysis of transcript data. They are not meant to be associated with individual transcripts, but rather with all data for a particular individual. The first weight is a simple adjustment for nonresponse to the transcript study itself, where response is defined as an eligible case having one or more coded transcript records in the data file. The other two adjusted weights account for multiple instances of nonresponse (e.g., no transcripts, no response to the fourth follow-up survey, missing data for critical items). Nonresponse

adjustments were computed as ratio adjustments within 39 separate weighting classes. Cases were assigned to each weight class based on sex, race/ethnicity, high school grades, and high school program, and within each group by whether or not only proprietary schools were attended. The final adjusted weights are the product of the raw weight for the “completed” case and the nonresponse adjustment factor for the weighting class to which the case belongs.

Imputation. The problem of missing data was resolved for certain items by supplemental data collections, the creation of composite variables, and some imputation of activity state and other variables. Most of the variables were created by pooling information from various items. For example, the activity states for 1972 and 1973 were updated with information gleaned from the Activity State Questionnaires that were administered concurrently with second follow-up operations. While some procedures for imputing missing data for activity state variables were incorporated in the steps of defining and recoding variables, two further phases of imputation procedures were implemented. The first phase involved direct logical inferences (e.g., type of school from name and address of school); the second phase involved indirect logical inferences (e.g., impute studying full time for those whose study time is unknown but who are studying and not working).

5. DATA QUALITY AND COMPARABILITY

The survey was implemented after an extensive period of planning, which included the design and field test of survey instrumentation and procedures. Any additional questions were field-tested prior to inclusion in the survey. The NLS:72 sample design and weighting procedures assured that participants’ responses could be generalized to the population of interest. Quality control activities were used throughout the data collection and processing of the survey.

Sampling Error

Statistical estimates derived from NLS:72 data are subject to sampling variability. Like almost all national samples, the NLS:72 sample is not a simple random sample. Taylor Series estimation techniques were used to compute standard errors in published NLS:72 reports.

In addition to standard errors, it is often useful to report design effects and the root mean design effect for complex surveys, such as NLS:72. Results from several NLS:72 studies suggest that a straightforward multiplicative adjustment of the simple random sample standard error equation adequately estimates the actual standard error estimate for a percentage. The three generalized mean design effects for the first, second, and third follow-up surveys are, respectively, the square root of 1.39, 1.35, and 1.44. To be conservative, the highest value—the square root of 1.44—can be used as an estimate for fourth follow-up data. For the fifth follow-up, the mean design effect for the overall NLS:72 sample is 2.64. The mean design effects indicate that an estimated percentage in the NLS:72 data is—on average—more than twice as variable as the corresponding statistic from a simple random sample of the same size. The mean design effects vary across the domains from a low of 2.0 for respondents from the highest SES quartile to a high of 3.8 for Black respondents.

Nonsampling Error

The major sources of nonsampling error in NLS:72 were coverage error and nonresponse error.

Coverage error. To identify public schools not included in the original sample frame, an additional 200 school districts were contacted after the base-year survey was completed, resulting in the identification of 45 augmentation schools. To compensate for the base-year undercoverage, samples of former 1972 high school seniors from 16 of these schools were included in the first and subsequent follow-up surveys. In addition, at the end of the base-year survey, several strata had no participating schools and many more had only one school (whereas the original sample design called for two schools). To compensate for this large school nonresponse, 205 base-year noncooperating primary schools and 36 backup schools were added to the sample prior to the first follow-up survey for “resurveying” with the original design. The former 1972 high school seniors from these augmented and resurveyed schools were asked some retrospective (senior year) questions during the first follow-up survey. These individuals—who redress the school frame undercoverage bias in the base year—do not appear in the NLS:72 base-year files that would typically be employed for comparisons of high school seniors; however, the presence of some retrospective data for these individuals permits refinement of comparisons grounded in 1972 data.

Also, while every effort was made to include in the fifth follow-up all persons with teaching experience, it is

conceivable that some individuals who entered teaching late were among the 6,000 cases not included in the fifth follow-up subsample. These individuals would not have had a chance to participate in the Teaching Supplement.

Nonresponse error. Detailed rates of response to various surveys and the availability of specific data items are provided in *NLS:72 Fifth Follow-Up (1986) Final Technical Report* (Sebring et al. 1987).

Unit nonresponse. For the NLS:72 student surveys, there were two stages of sample selection and hence two types of unit nonresponse—school and student. During the base year, sample schools were asked to permit the selection of individual high school seniors for the collection of questionnaire and test data. Schools that refused to cooperate in either stage of sample selection were dropped from the sample. The bias introduced by base-year school-level refusals is of particular concern since it carried over into successive rounds of the survey. To the extent that the students in refusal schools differed from students in cooperating schools during later survey waves, the bias introduced by base-year school nonresponse persisted from one wave to the next. (Base-year school nonresponse is addressed under “Coverage error” above.)

Also, individual students at cooperating schools could fail to take part in the base-year survey. Student nonresponse would not necessarily carry over into subsequent waves since student nonrespondents in the base year remained eligible for sampling throughout the study. However, a study of third follow-up responses indicated that response to earlier survey waves was the most important predictor of response to the third follow-up.

Due to intensive data collection procedures, the response rates to the individual NLS:72 surveys were high (80 percent or better) among eligible sample members. At the conclusion of fourth follow-up activities, a total of 12,980 individuals had provided information in each of the first five survey waves (base-year and all four follow-up surveys), representing 78 percent of the 16,680 base-year respondents. As a result of the various retrospective data collection efforts, the number of individuals with some key data elements for all time points through the fourth follow-up survey is 16,450—73 percent of the 22,650 respondents who participated in at least one survey. In conjunction with the supplemental data collection efforts, this led to a high degree of sample integrity among the key longitudinal data elements.

Only sample members who had participated in at least one of the previous five waves were eligible for selection into the fifth follow-up sample. Of the 14,430 fifth follow-up sample members (excluding the

deceased), 89.0 percent (unweighted) completed questionnaires in the fifth follow-up; 92.2 percent participated in at least five of the six waves; and 62.1 percent participated in all six waves. There was moderate variation in weighted nonresponse rates by region; nonresponse was greater in the West and Northeast, lower in the South, and lowest in the North Central region. The relationship between urbanization and nonresponse was about the same as for region—13 percent for rural schools, 15 percent for urban schools, and 18 percent for suburban schools. There was marked variation in nonresponse by race; Blacks showed the highest nonresponse (22.1 percent), followed closely by Hispanics (19.8 percent) and Whites (14.0 percent). Males had a higher nonresponse rate (17.3 percent) than females (13.6 percent).

In PETS, one or more transcripts were received for 91.1 percent of the 13,830 sample members reporting postsecondary school attendance since leaving high school. A single transcript was received for 55 percent of this group, two transcripts for 27 percent, and three or more transcripts for over 9 percent. At the transcript level, 87 percent of the 21,870 “in-scope” transcripts requested were supplied by the postsecondary schools (2,570 of the 24,430 transcripts initially requested could not be obtained because the school had no record of the student’s attendance). Response rates varied from a high of 93 percent for transcripts sought from public 4-year colleges and universities to a low of 55 percent from vocational and proprietary schools. The higher response rates for public and private nonvocational schools may be attributable to their typically longer period of existence and the relative permanence of their student files. Telephone follow-up calls to nonresponding schools revealed that nearly half of the vocational school transcripts requested for NLS:72 students were unavailable.

Item nonresponse. While unit nonresponse can be adjusted for by weighting, this approach is impractical for item nonresponse. Researchers should take into account that NLS:72 respondents often skipped questions incorrectly or gave unrecognizable answers. However, efforts were made to retrieve missing data for critical items by telephone, with a success rate of over 90 percent.

Most item nonresponse in NLS:72 resulted from respondents’ limited recall of past events or misinterpretation of questions and routing instructions. Many items in the student files appear to have high nonresponse rates (i.e., above 10 percent). In most instances, these items are associated with the routing, or skip, patterns in the instruments. (A routing question is one that implicitly or explicitly directs a respondent

around other questions in the instrument.) Rather conservative rules were used to label blanks as either missing (illegitimate skip—code 98) or inapplicable (legitimate skip—code 99). With the more complex routing patterns, a large section of items was sometimes coded illegitimate (code 98) due to just one inconsistency in the pattern. The data user should be careful in interpreting data coded 98 and 99 and should further examine data that lie within complex routing patterns when they are required for analysis. Similarly, data labeled as suspect during the editing stage should be reexamined and possibly reclassified for specific analytic purposes.

Measurement error. The survey data were monitored for quality of processing and evaluated to determine the extent of any problems and the sources of errors. Some examples are given below.

Study of edit failures. If the respondent failed to answer certain key items properly, the questionnaire failed an edit and the respondent was contacted by telephone. A special study of survey responses in the third follow-up was conducted to determine why so many questionnaires (over 60 percent) failed the edit process. This study concluded that (1) the majority of edit failures associated with itemized financial questions involved the respondent's failure to supply answers to each of the requested line items; (2) items structured as "check all responses that apply" were likely to be failed by a substantial number of respondents; and (3) overall data entry errors were low (except for items requiring itemized financial information).

Review of routing patterns. Quality control, completeness, routing, and consistency indices were created for use with the student files. Routing indices, computed identically for each survey, indicate the percentage of the routing questions that were ambiguously answered by an individual for a given instrument. The first four follow-up questionnaires contained 33, 52, 67, and 61 routine patterns, respectively. In general, 56 to 68 percent of all respondents proceeded through an instrument without violating any routing patterns; about 20 to 30 percent violated 1 to 5 routing patterns; and 7 to 15 percent violated 6 to 10 patterns. In all four instruments, a small percentage (3 to 7 percent) of sample members had great difficulty with the routing patterns and violated the instructions in more than 10 different patterns.

Monitoring of data entry. For the first four follow-up surveys, direct data entry terminals were used to key the survey data. For the Supplemental Questionnaires

administered in the fourth follow-up survey, data entry error rates were computed based on three keyings. After the initial keying, a random sample of the questionnaires from each batch was selected for rekeying by two additional operators. The results were within the overall error rate tolerance established for NLS:72. The variable error rate across samples and operators on the selected questionnaires was 0.00040; the estimated character error rate was 0.00023.

Data Comparability

One of the major goals of the NLS Program is to make the data sufficiently comparable to allow cross-cohort comparisons between studies (NLS:72 vs. HS&B vs. NELS:88 vs. ELS:2002), as well as comparative analyses of data across waves of the same study. Nevertheless, data users should be aware of some variations in sample design, questionnaire and test content, and data collection methods that could impact the drawing of valid comparisons.

Sample design changes. Although the general NLS:72 sample design was similar for all waves, there are some differences worth noting. The original sample design called for two schools to be surveyed from each of 600 strata; however, at the end of the base-year survey, several strata had no participants and many more had only one. As a result of a resurvey effort during the first follow-up survey, the final sample included *at least* two participating schools from each stratum. The fifth follow-up sample design differed from the base-year design in that the student selection probabilities were equal in the base-year design but unequal in the fifth follow-up.

Reporting period differences. The first four follow-ups requested data as of October of the survey year, whereas the fifth follow-up used February 1986 as the reference date.

Content changes. Due to the increased interest in event history analysis, the fifth follow-up survey collected more detailed information than did earlier surveys on the time periods during which respondents held jobs or were in school. Instead of recording one start and stop date for each school and job, up to eight time periods (or start and stop dates) were shown. To allow for maximum user flexibility, the responses were coded into pairs of start and stop dates.

Comparisons between NLS:72 student data and PETS data. There are substantial discrepancies between student-reported postsecondary attendance in the NLS:72 follow-up surveys and the evidence obtained from official school transcripts collected in the PETS. One interpretation is that NLS:72 respondents

overreported instances of postsecondary school attendance by about 10 percent (unweighted). If so, researchers analyzing postsecondary schooling using only the survey data would overestimate significantly the extent of this activity. Coding errors could offer further explanation for the discrepancies.

Comparisons among NLS:72, HS&B, NELLS:88, and ELS:2002. The four NLS studies were specifically designed to facilitate comparisons with each other. At the student level, three different kinds of comparative analyses are possible. (See Section 2. Uses of Data for more detail.) The overall sample design is similar, and a core of questionnaire items is comparable across all four studies. Additionally, item response theory methods can be used to place mathematics, vocabulary, and reading scores on the same scale for 1972, 1980, 1992, and 2004 high school seniors.

However, despite the considerable similarities among NLS:72, HS&B, NELLS:88, and ELS:2002, the differences in sample definition and statistical design have implications for intercohort analysis. Also, sampling error tends to be a greater problem for intercohort comparisons than for intracohort comparisons because there is sampling error each time an independent sample is drawn. In addition, a number of nonsampling errors may arise when estimating trends based on results from two or more sample surveys. For example, student response rates differ across the four NLS studies, and the characteristics of the nonrespondents may differ as well. The accuracy of intercohort comparisons may also be influenced by differences in context and question order for trend items in the various student questionnaires; differences in test format, content, and context; and other factors, such as differences in data collection and methodology. While some effort has been made to maintain trend items over time, strict test and questionnaire overlap is not considerable across the four NLS studies. More specifically, differences exist in questionnaire construction and in mode and type of survey administration. See chapter 7 (HS&B), chapter 8 (NELLS:88), and chapter 9 (ELS:2002) for additional information on the comparability of the four NLS studies.

6. CONTACT INFORMATION

For content information on NLS:72, contact:

Aurora D'Amico
Phone: (202) 502-7334
E-mail: aurora.damico@ed.gov

Mailing Address:

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5652

7. METHODOLOGY AND EVALUATION REPORTS

General

Jones, C. (1985). *The National Longitudinal Study of the High School Class of 1972 (NLS-72) Fifth Follow-Up Survey and High School and Beyond Third Follow-Up Survey: Field Test Report* (ED 269-465). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

McAdams, K. (1981). *National Longitudinal Study of the High School Class of 1972: An Historical Overview and Summary* (ED 217-051). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Quinn, P. (1995). *National Education Longitudinal Study of 1988: Second Follow-up Research and Development Working Papers* (NCES 94-251). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Riccobono, J.A. (1981). *National Longitudinal Study of the High School Class of 1972: Fourth Follow-Up Survey Final Methodological Report* (ED 217-052). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Sebring, P., Tourangeau, R., Spencer, B., Glusberg, M., Campbell, B., Singleton, M. (1987). *National Longitudinal Study of the High School Class of 1972 (NLS-72) Fifth Follow-Up (1986) Final Technical Report* (NCES 87-100). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Uses of Data

Ingels, S., and Baldrige, J. (1995). *National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELLS:88 Seniors* (NCES Working Paper 95-05). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Survey Design

Rock, D.A., Hilton, T.L., Pollack, J.M., Ekstrom, R.B., and Goertz, M.E. (1985). *Psychometric Analysis of the NLS-72 and the High School and Beyond Test Batteries* (NCES 85-218). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Spencer, B., Sebring, P., and Campbell, B. (1987). *The National Longitudinal Study of the High School Class of 1972 (NLS-72) Fifth Follow-up (1986) Sample Design Report* (CS 88-403c). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Westat. (1972). *Sample Design for the Selection of a Sample of Schools with Twelfth-Graders for a Longitudinal Study*. National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Cahalan, M.W., Ingels, S.J., Burns, L.J., Planty, M., and Daniel, B. (2006). *United States High School Sophomores: A Twenty-Two Year Comparison, 1980–2002* (NCES 2006-327). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Green, P., Dugoni, B.L., and Ingels, S.J. (1995). *Trends Among High School Seniors, 1972–1992* (NCES 95-380). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Henderson, L.B., and Allen, D.R. (1981). *NLS Data Entry Quality Control: The Fourth Follow-Up Survey* (ED 221-593). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Williams, S.R., and Folsom, R.E. (1977). *Bias Resulting From School Nonresponse: Methodology and Findings*. Washington, DC.

Wisembaker, J.M. (1981). *Factors Associated With Edit Failure* (NCES 82-213). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Wisembaker, J.M., and Kolstad, A.J. (1981.) *Factors Related to Third Follow-Up Survey Responses* (NCES 82-209). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Chapter 7: High School and Beyond (HS&B) Longitudinal Study

1. OVERVIEW

The High School and Beyond (HS&B) Longitudinal Study was the second study conducted as part of NCES' National Longitudinal Studies Program. This program was established to study the educational, vocational, and personal development of young people, beginning with their elementary or high school years and following them over time as they take on adult roles and responsibilities. The HS&B included two high school cohorts—a senior cohort (the graduating class of 1980) and a sophomore cohort (the sophomore class of 1980). Students, school administrators, teachers, parents, and administrative records provided data for the study. HS&B results can be compared with the results of three other longitudinal studies—the National Longitudinal Study of the High School Class of 1972 (NLS:72), the National Education Longitudinal Study of 1988 (NELS:88), and the Education Longitudinal Study of 2002 (ELS:2002). (See chapters 6, 8, and 9, respectively, for descriptions of these studies.)

The HS&B covered more than 30,000 high school seniors and 28,000 high school sophomores. It consisted primarily of a base-year survey in 1980 and four follow-up surveys in 1982, 1984, 1986, and 1992. Record studies were also conducted to obtain key supplemental data on students. As part of the first follow-up, high school transcripts were requested for the sophomore cohort, providing information on the sophomores' course taking behavior through their 4 years of high school. Postsecondary transcripts were collected in 1984 for the senior cohort and in 1987 and 1993 for the sophomore cohort. In addition, student financial aid data were obtained from administrative records in 1984 for the senior cohort and in 1986 for the sophomore cohort. The HS&B project ended in 1993 after the completion of the fourth follow-up survey and a related transcripts study of the sophomore cohort.

Purpose

To (1) study longitudinally the given cohorts' educational, vocational, and personal development, beginning with their high school years, and the personal, familial, social, institutional, and cultural factors that may affect that development; and (2) compare the results with data from the NLS:72, NELS:88, and ELS:2002 to facilitate cross-cohort studies of American youth's schooling and socialization.

Components

The HS&B compiled data from a sample of students, parents, teachers, and school administrators in a base-year and four follow-up surveys. It also collected high school and postsecondary transcripts and administrative financial aid records. The various components are described below.

LONGITUDINAL SAMPLE SURVEY OF THE HIGH SCHOOL SOPHOMORE AND SENIOR CLASSES OF 1980; BASE- YEAR SURVEY AND FOUR FOLLOW-UPS, ENDING IN 1992

HS&B collects data from:

- Students and dropouts
- School administrators
- Teachers
- Parents
- High school transcripts
- Postsecondary transcripts
- Postsecondary financial aid records

Base-Year Survey. The base-year survey was conducted in spring 1980 and comprised the following:

Student Questionnaire. Students were asked to (1) fill out a booklet, which included several items on the use of non-English languages as well as confidential identifying information; (2) complete a questionnaire that focused on their individual and family background, high school experiences, work experiences, future educational plans, future occupational goals, and plans for and ability to finance postsecondary education; and (3) take timed cognitive tests that measured verbal and quantitative abilities. The sophomore test battery included achievement measures in science, writing, and civics, while seniors were asked to respond to tests measuring abstract and nonverbal abilities.

School Questionnaire. Completed by an official in the participating school, this questionnaire collected information about enrollment, staff, educational programs, facilities and services, dropout rates, and special programs for handicapped and disadvantaged students.

Teacher Comment Checklist. At each grade level, teachers had the opportunity to answer questions about the traits and behaviors of sampled students who had been in their classes. The typical student in the sample was rated on average by four different teachers.

Parent Questionnaire. A sample of parents provided information about family attitudes, family income, employment, occupation, salary, financial planning, and how these affect postsecondary education and goals. The results included responses from the parents of about 3,600 sophomores and 3,600 seniors.

First Follow-up Survey. The first follow-up survey was conducted in spring 1982. As in the base-year survey, information was collected from students, school administrators, and parents. For the 1980 senior cohort, high school and postsecondary experiences were the main focus of the survey; seniors were asked about their school and employment experiences, family status, and attitudes and plans. For the 1980 sophomore cohort, the survey gathered information on school, family, work experiences, educational and occupational aspirations, personal values, and test scores of sample participants. A high school transcript collection was also part of the first follow-up for sophomore cohort members. (See section on Record Studies for more detail.)

Sophomores were classified by high school status as of 1982 (i.e., dropout, same school, transfer, or early graduate). Dropouts completed a *Not Currently in High*

School Questionnaire, which included some questions from the regular Student Questionnaire but focused on their reasons for dropping out and its impact on their educational and career development. In addition to the regular Student Questionnaire, a *Transfer Supplement* was completed by members of the sophomore cohort who had transferred out of their base-year sample high school to another high school. This supplement gathered information on the reasons for transferring and for selecting a particular school, the length of the interruption in schooling and why it occurred, and particulars about the school itself (type, location, entrance requirements, size of student body, grades). Sophomore cohort members who graduated from high school ahead of schedule completed an *Early Graduate Supplement* in addition to the regular questionnaire. The Early Graduate Supplement documented the reasons for and circumstances of early graduation, the adjustments required to finish early, and respondents' activities compared with those of other out-of-school survey members (i.e., dropouts, 1980 seniors).

Second Follow-up Survey. This survey was conducted in spring 1984. For both the sophomore and senior cohorts, the survey collected data on the students' work experience, postsecondary schooling, earnings, periods of unemployment, and so forth. For seniors, postsecondary transcripts and financial aid records were also collected. (See section on record studies for more detail.)

Third Follow-up Survey. This survey was administered in spring 1986, using the same questionnaire for both the sophomore and senior cohorts. To maintain comparability with prior waves, many questions from earlier follow-up surveys were repeated. Respondents were asked to update background information and to provide information about their work experience, unemployment history, education and other training, family information (including marriage patterns), income, and other experiences and opinions. Financial aid records and postsecondary transcripts were collected for sophomores. (See section on Record Studies for more detail.)

Fourth Follow-up Survey. This survey was administered in spring 1992 only to the sophomore cohort. The survey sought to obtain valuable information on issues of access to, and choice of, undergraduate and graduate education institutions; persistence in obtaining educational goals; progress through the curriculum; rates of degree attainment and other assessments of educational outcomes; and rates of return to the individual and society. Additionally, a collection of postsecondary transcripts for sophomore

cohort members (i.e. members who had received their baccalaureate degrees and then went on to pursue graduate, doctoral, and first-professional degrees) took place in 1993.

Record Studies. The following record studies were conducted during the course of the HS&B project.

High School Transcript Study. In fall 1982, as part of the first follow-up, nearly 16,000 high school transcripts were collected for sophomore cohort students who were seniors in 1982. This data collection allowed the study of the course taking behavior of the members of the sophomore cohort throughout their 4 years of high school. Data included a six-digit course number for each course taken; course credit, expressed in Carnegie units (a standard of measurement that represents one credit for the completion of a 1-year course); course grade; year that course was taken; grade point average; days absent; and standardized test scores. (For more information, see chapter 29, which covers the High School Transcript Studies.)

Postsecondary Education Transcript Study. This study gathered data on students' academic histories since leaving high school. As part of the second follow-up in 1984, postsecondary transcripts were collected for the senior cohort. Transcripts were requested from all postsecondary institutions reported by senior cohort members in the first and second follow-up surveys. Transcript data included dates of attendance; fields of study; degrees earned; and the titles, grades, and credits of every course attempted at each institution.

In 1987 and again in 1993, postsecondary transcripts were collected for the sophomore cohort. The latter collection allowed information to be obtained on sophomore cohort members who had received their baccalaureate degrees and then went on to pursue graduate, doctoral, and first-professional degrees.

Student Financial Aid Records. In 1984, HS&B collected institutional financial aid records and federal records on the Guaranteed Student Loan Program and the Pell Grant Program for seniors who had indicated postsecondary attendance. Federal financial aid records were obtained for the sophomore cohort in 1986.

Periodicity

The base-year survey was conducted in 1980, with four follow-ups in 1982, 1984, 1986, and 1992 (the 1992 follow-up included only the sophomore cohort). High school transcripts were collected for the sophomore cohort in 1982. Postsecondary transcripts were collected for the senior cohort in 1984 and for the sophomore cohort in 1987 and 1993. Student financial

aid records were collected for the senior cohort in 1984 and for the sophomore cohort in 1986.

2. USES OF DATA

The HS&B provides information on the educational, vocational, and personal development of young people as they move from high school into postsecondary education or the workforce and then into adult life. The initial longitudinal study (NLS:72) laid the groundwork for comparison with HS&B, while successive studies (NELS:88 and ELS:2002) provide a basis for further comparisons. NLS:72 recorded the economic and social conditions surrounding high school seniors in 1972 and, within that context, their hopes and plans; subsequently, it measured outcomes while also observing the intervening processes. Data on 1980 seniors from the HS&B base-year survey are directly comparable to NLS:72 data on 1972 seniors. With the follow-up data, trend comparisons can be made for the period 1972 to 1984. HS&B permits researchers to further monitor change by, for example, measuring the economic returns of postsecondary education for minorities and delineating the need for financial aid.

By following adolescents at an earlier age (beginning in eighth grade) and into the 21st century, NELS:88 expands the base of knowledge established in the NLS:72 and HS&B studies. NELS:88 first follow-up data provide a comparison point to high school sophomores 10 years earlier, as studied in HS&B; the second follow-up data allow trend comparisons of the high school class of 1992 with the 1980 seniors studied in the HS&B. The third follow-up allows comparisons with HS&B related to postsecondary outcomes. (Please see chapter 8 for detailed information on NELS:88.)

ELS:2002 further measures 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. Comparisons can be made between high school sophomores in 1980 and in 2002, and between high school seniors in 1980 and in 2004 (the first follow-up of ELS:2002) using the HS&B and ELS:2002 studies. (Please see chapter 9 for detailed information on ELS:2002.)

By comparing the results of the HS&B and its three related longitudinal studies, researchers can determine how plans and outcomes differ in response to changing conditions, or remain the same despite such changes.

The HS&B allows both cross-sectional and longitudinal analyses of the students who were sophomores or seniors in 1980. The data are used to address issues of educational attainment, employment, family formation, personal values, and community activities since 1980. For example, a major study on high school dropouts used HS&B data to demonstrate that a large number of dropouts return to school and earn a high school diploma or an equivalency certificate. Other examples of issues and questions that can be addressed are as follows:

- How, when, and why do students enroll in postsecondary education institutions?
- Do students who, while in high school, expect to complete the baccalaureate degree actually do so?
- How has the percentage of recent graduates from a given cohort who enter the workforce in their field changed over the past years?
- What are the long-term effects of not completing high school in the traditional way? How do employment and earnings event histories of traditional high school graduates differ from those of students who do not finish high school in the traditional manner?
- Do individuals who attend college earn more than those who do not attend college? What is the effect of student financial aid?
- What percentage of college graduates is eligible or qualified to enter a public service profession, such as teaching?
- How many college graduates enter the workforce full time in the area for which they are qualified?
- How, and in what ways, do public and private schools differ?

3. KEY CONCEPTS

Some of the key terms related to HS&B are defined below.

Cognitive Tests. Achievement tests administered to both cohorts in the base-year survey and only to the sophomore cohort in the first follow-up. For the

sophomore cohort, the content in the base-year and first follow-up achievement tests was as follows: (1) vocabulary (21 items, 7 minutes), using a synonym format; (2) reading (20 items, 15 minutes), consisting of short passages (100–200 words) followed by comprehension questions and a few analysis and interpretation items; (3) mathematics (38 items, 21 minutes), in which students were asked to determine which of two quantities was greater, whether they were equal, or whether there were insufficient data to answer the question; (4) science (20 items, 10 minutes), based on science knowledge and scientific reasoning ability; (5) writing (17 items, 10 minutes), based on writing ability and knowledge of basic grammar; and (6) civics education (10 questions, 5 minutes), based on various principles of law, government, and social behavior. Seniors in the base-year survey were given a cognitive test with items in the following categories: vocabulary (27 items, 9 minutes), reading (20 items, 15 minutes), mathematics (33 items, 19 minutes), picture-number pairs (15 items, 5 minutes), mosaic comparisons (89 items, 6 minutes), visualization in three dimensions (16 items, 9 minutes), and questions about the test (5 minutes).

Course Offering and Course Taking. Course offering data were collected from the School Questionnaires filled out by school administrators; course offerings included regular and advanced placement curricula provided by the schools. Course taking data were collected in different ways for the sophomore and senior cohorts. For sophomores, official high school transcripts provided records of students' coursework. For the senior cohort, high school transcripts were not available; instead, coursework was self-reported by seniors in a series of items asking retrospectively about the courses and hours taken. Despite these differences in data collection, the listings of courses for the two cohorts were consistent, including major subjects in both regular and advanced placement curricula.

Socioeconomic Status (SES). The level of a student's SES was a composite variable, constructed from a set of variables from the base-year and first follow-up data, including father's occupation, father's education, mother's education, family income, and material possessions in the household.

4. SURVEY DESIGN

Target Population

High school students who were in the 10th or 12th grade in U.S. public and private schools in spring 1980.

Sample Design

HS&B was designed to provide nationally representative data on 10th- and 12th-grade students in the United States.

Base-Year Survey. In the base-year, students were selected using a two-stage, stratified probability sample design, with secondary schools as the first-stage units and students within schools as the second-stage units. Sampling rates were set so as to select in each stratum the number of schools needed to satisfy study design criteria regarding minimum sample sizes for certain types of schools. The following types of schools were oversampled to make the study more useful for policy analyses: public schools with a high percentage of Hispanic students; Catholic schools with a high percentage of Black, Hispanic, and other race/ethnicity students; alternative public schools; and private schools with high-achieving students. Thus, some schools had a high probability of inclusion in the sample (in some cases, equal to 1.0), while others had a low probability. The total number of schools in the sample was 1,120, selected from a frame of 24,730 schools with grades 10 or 12 or both (there was only one school sample in the base-year for both cohorts). Within each stratum, schools were selected with probabilities proportional to the estimated enrollment in their 10th and 12th grades.

Within each school, 36 seniors and 36 sophomores were randomly selected. In schools with fewer than 36 seniors or 36 sophomores, all eligible students were drawn in the sample. Students in all but the special strata were selected with approximately equal probabilities. (The students in the special strata were selected with higher probabilities.) Special efforts were made to identify sampled students who were twins or triplets so that their co-twins or co-triplets could be invited to participate in the study.

Substitution was carried out for schools that refused to participate in the survey. There was no substitution for students who refused, for students whose parents refused, or for students who were absent on survey day and makeup days.

First Follow-up Survey. The first follow-up sophomore and senior cohort samples were based on the base-year samples, retaining the essential features of a stratified multistage design. (For details see *High School and Beyond First Follow-Up (1982) Sample Design Report* [Tourangeau et al. 1983].)

For the sophomore cohort, all schools selected for the base-year sample were included in the first follow-up (except 40 schools that had no 1980 sophomores, had closed, or had merged with other schools in the

sample). The sample also included 17 schools that received two or more students from base-year schools; school-level data from these institutions were eventually added to students' records as contextual information. However, these schools were not added to the existing probability sample of schools.

Sophomores still enrolled in their original base-year schools were retained with certainty since the base-year clustered design made it relatively inexpensive to resurvey and retest them. Sophomores no longer attending their original base-year schools were subsampled (i.e., dropouts, early graduates, students who transferred as individuals to a new school). Certain groups were retained with higher probabilities in order to support statistical research on such policy issues as excellence of education throughout society, access to postsecondary education, and transition from school to the labor force.

Students who transferred as a class to a different school were considered to be still enrolled if their original school had been a junior high school, had closed, or had merged with another school. Students who had graduated early or had transferred as individuals to other schools were treated as school leavers for the purposes of sampling. The 1980 sophomore cohort school leavers were selected with certainty or according to predesignated rates designed to produce approximately the number of completed cases needed for each of several different sample categories. School leavers who did not participate in the base-year were given a selection probability of 0.1.

For the 1980 senior cohort, students selected for the base-year sample had a known, nonzero chance of being selected for the first and all subsequent follow-up surveys. The first follow-up sample consisted of 11,995 selections from the base-year probability sample (including 11,500 of the 28,240 base-year participants and 495 of the 6,740 base-year nonparticipants). In addition, 204 nonsampled co-twins or co-triplets (who were not part of the probability sample) were included in the first follow-up sample, resulting in a total of 12,200 selections.

High School Transcript Study (1980 Sophomore Cohort). Subsequent to the first follow-up survey, high school transcripts were sought for a probability subsample of nearly 18,500 members of the 1980 sophomore cohort. The subsampling plan for the transcript study emphasized the retention of members of subgroups of special relevance for education policy analysis. Compared to the base-year and first follow-up surveys, the transcript study sample design further increased the overrepresentation of certain

race/ethnicity groups, students who attended private high schools, school dropouts, transfers, early graduates, and students whose parents completed the base-year Parent Questionnaire on financing postsecondary education. Transcripts were collected and processed for nearly 16,000 members of the sophomore cohort.

Second and Third Follow-up Surveys. The sample for the second follow-up survey of the 1980 sophomore cohort was based upon the design of the High School Transcript Study. A total of 14,830 cases were selected from the nearly 18,500 sample members retained for the transcript study. The second follow-up sample included disproportionate numbers of sample members from policy-relevant subpopulations. The sample for the senior cohort in the second follow-up consisted exactly of those sample members selected into the first follow-up sample. The senior and sophomore cohort samples for the third follow-up survey were the same as those used for the second follow-up. The third follow-up was the last survey conducted for the senior cohort. Postsecondary school transcripts were collected for all members of the senior cohort who reported attending any form of postsecondary schooling in either of the follow-up surveys. Over 7,000 individuals reported more than 11,000 instances of postsecondary school attendance.

Fourth Follow-up Survey. The fourth follow-up was composed solely of members of the sophomore cohort, and consisted exactly of those students selected into the second and third follow-up sample. For any student who had ever enrolled in postsecondary education, complete transcript information was requested from the institutions indicated by the student.

Data Collection and Processing

HS&B compiled data from six primary sources: students, school administrators, teachers, parents of selected students, high school administrative records (transcripts), and postsecondary administrative records (transcripts and financial aid). Data collection began in fall 1979 (when information from school administrators and teachers was first gathered) and ended in 1993 (when postsecondary transcripts of sophomore cohort members were collected). The National Opinion Research Center (NORC) at the University of Chicago was the contractor for the HS&B project.

Reference dates. In the base-year survey, most questions referred to the students' experience up to the time of the survey administration in spring 1980 (i.e., all 4 high school years for the senior cohort and the first 2 high school years for the sophomore cohort). In

the follow-ups, most questions referred to experiences that occurred between the previous survey and the current survey. For example, the second follow-up largely covered the period between 1982 (when the first follow-up was conducted) and 1984 (when the second follow-up was conducted).

Data collection. In both the base-year and first follow-up surveys, it was necessary to secure a commitment to participate in the study from the administrator of each sampled school. For public schools, the process began by contacting the chief state school officer. Once approval was gained at the state level, contact was made with district superintendents and then with school principals. Wherever private schools were organized into an administrative hierarchy (e.g., catholic school dioceses), approval was obtained at the superior level before approaching the school principal or headmaster. The principal of each cooperating school designated a school coordinator to serve as a liaison between the NORC staff, school administrator, and selected students. The school coordinator (most often a senior guidance counselor) handled all requests for data and materials, as well as all logistical arrangements for student-level data collection on the school premises.

In the 1980 base-year survey, a single data collection method—on-campus administration—was used for both the sophomore and senior cohorts. In the first follow-up, most members of the sophomore cohort (nearly all of whom were then in the 12th grade) were resurveyed using methods similar to those of the base-year survey. However, since some of the 1980 sophomores had left school by 1982, the first follow-up survey involved on-campus administration for in-school respondents as well as off-campus group administration for school leavers (transfers, dropouts, early graduates). On-campus surveys generally were similar to those used in the base-year. Off-campus survey sessions were held afterward for school leavers in the sophomore cohort. Personal or telephone interviews were conducted with individuals who did not attend the sessions. Members of the 1980 senior cohort were surveyed primarily by mail. Nonrespondents to the mail survey (approximately 25 percent) were interviewed either in person or by telephone.

By the time of the second follow-up, the sophomore cohort was out of school. Thus, in the second (1984) and third (1986) follow-ups, data for both the sophomore and senior cohorts were collected through mailed questionnaires. Telephone and personal interviews were conducted with sample members who did not respond to the mailed survey within 2 to 3 months. Only the sophomore cohort was surveyed in

the fourth follow-up (1992). Computer-assisted telephone interviewing (CATI) was used to collect these data. The CATI program included two instruments; the first was used to locate and verify the identity of the respondent, while the second contained all of the survey questions. The average administration time for an interview was 30.6 minutes. Intensive telephone locating and field intervention procedures were used to locate respondents and conduct interviews.

Processing. Although procedures varied across survey waves, all Student Questionnaires in all waves were checked for missing critical items. Approximately 40 items in each of the main survey instruments were designated as critical or “key” items. Cases failed this edit, if a codable response was missing for any of the key items. Such cases were flagged and then routed to the data retrieval station, where staff called respondents to obtain missing information or otherwise resolve the edit failure.

The base-year procedures for data control and preparation differed significantly from those in the follow-up surveys. Since the base-year student instruments were less complex than later instruments, the completed documents were sent directly from the schools to NORC’s optical scanning subcontractor for conversion to machine-readable form. The scanning computer was programmed to perform the critical item edit on Student Questionnaires and to generate listings of cases missing critical data, which were then sent to NORC for data retrieval. School and Parent Questionnaires were converted to machine-readable form by the conventional key-to-disk method at NORC.

All follow-up questionnaires were sent to NORC for receipt control and data preparation prior to being shipped to the scanning subcontractor. The second follow-up survey contained optically scannable grids for the answers to numeric questions; staff examined numeric responses for correct entry (e.g., right justification, omission of decimal points). In the third follow-up, a portion of the instrument was designed for computer-assisted data entry (CADE), while the rest was prepared for optical scanning. All major skip items and all critical items were entered by CADE. With this system, operators were able to combine data entry with the traditional editing procedures. The CADE system stepped question by question through critical and numeric items, skipping over questions that were slated for scanning and questions that were legitimately skipped because of a response to a filter question. Ranges were set for each question, preventing the accidental entry of illegitimate responses. CADE

operators were also responsible for the critical item edit; those critical items that did not pass the edit were flagged for retrieval, both manually and by the CADE system. After the retrieved data were keyed, questionnaires were shipped to the scanning firm.

For the fourth follow-up, a CATI program captured the data at the time of the interview. The CATI program examined the responses to completed questions and used that information to route the interviewer to the next appropriate question. It also applied the customary edits, described below under “Editing.” At the conclusion of an interview, the completed case was deposited in the database ready for analysis. There was minimal post-data entry cleaning because the interviewing module itself conducted the majority of necessary edit checking and conversion functions. A CADE system was designed to enter and code transcript data.

The first through fourth follow-ups required coding of open-ended responses on occupation and industry; postsecondary schools; major field of study for each postsecondary school; licenses, certificates, and other diplomas received; and military specialized schools, specialty, and pay grade. Coding was compatible with the coding done in NLS:72, using the same sources from NCES and the U.S. Bureau of the Census. (See chapter 6.) In the first follow-up, staff also coded open-ended questions in the Early Graduate and Transfer supplements, and transformed numeric responses to darkened ovals to facilitate optical scanning. In the third follow-up, all codes were loaded into a computer program for more efficient access. Coders typed in a given response, and the program displayed the corresponding numeric code.

In the fourth follow-up, interviewers received additional coding capabilities by temporarily exiting the CATI program and executing separate programs that assisted them in coding the open-ended responses. Data from the coding programs were automatically sent to the CATI program for inclusion in the dataset. In addition to the online coding tasks, interviewers recorded verbatim descriptions of industry and occupation. The coding scheme for industry in the fourth follow-up was a simplified version of the scheme used in previous rounds of HS&B (verbatim responses are available for more detailed coding). The coding scheme for occupation was adapted from verbatim responses received in the third follow-up. Postsecondary institutions were coded with Federal Interagency Committee on Education (FICE) codes.

Editing. In addition to the critical item edit described above, a series of edits checked the data for out-of-

range values and inconsistencies between related items. In the base-year, machine editing was limited to examining responses for out-of-range values. No interim consistency checks were performed since there was only one skip pattern.

In the first and second follow-ups, several sections of the questionnaire required respondents to follow skip instructions. Computer edits were performed to resolve inconsistencies between filter and dependent questions, detect illegal codes, and generate reports on the incidence of correctly and incorrectly answered questions. After improperly answered questions were converted to blanks, the student data were passed to another program for conversion to appropriate missing-data codes (e.g., “legitimate skip,” “refused”). Detection of out-of-range codes was completed during scanning for all questions except those permitting an open-ended response. Hand-coded data for open-ended questions (occupation, industry, institution, field of study) were matched by computer against lists of valid codes.

In the third follow-up, CADE carried out many of the steps that normally occur during machine editing. The system enforced skip patterns, range checking, and appropriate use of reserved codes—allowing operators to deal with problems or inconsistencies while they had the document in hand. For scanned items, the same machine-editing steps as those used in prior follow-ups were implemented. Since most of the filter questions were CADE-designated items, there were few filter-dependent inconsistencies to be handled in machine editing.

In the fourth follow-up, machine editing was replaced by the interactive edit capabilities of the CATI program, which tested responses for valid ranges, data field size, data type (numeric or text), and consistency with other answers or data from previous rounds. If the system detected an inconsistency due to a keying error by the interviewer, or if the respondent simply realized that he or she had made a reporting error earlier in the interview, the interviewer could go back and change the earlier response. As the new response was entered, all of the edit checks performed at the first response were again performed. The system then worked its way forward through the questionnaire using the new value in all skip instructions, consistency checks, and the like until it reached the first unanswered question, and control was then returned to the interviewer. When problems were encountered, the system could suggest prompts for the interviewer to use in eliciting a better or more complete answer.

Estimation Methods

Weighting is used to adjust for sampling and unit nonresponse.

Weighting. The weights are based on the inverse of the selection probabilities at each stage of the sample selection process and on nonresponse adjustment factors computed within weighting cells. While each wave provided weights for statistical estimation, the fourth follow-up weights can illustrate the concept of weighting. The fourth follow-up generated survey data and postsecondary transcript data. Weights were computed to account for nonresponse in both of these data collections.

First, a raw weight, unadjusted for nonresponse in any of the surveys, was calculated and included in the data file. The raw weight provided the basis for analysts to construct additional weights adjusted for the presence of any combination of data elements. However, *caution should be used if the combination of data elements results in a sample with a high proportion of missing cases*. For the survey data, two weights were computed. The first weight was computed for all fourth follow-up respondents. The second weight was computed for all fourth follow-up respondents who also participated in the base-year survey and in the first, second, and third follow-up surveys.

Two additional weights were computed to facilitate the use of the postsecondary transcript data. The collection of transcripts was based upon sophomore cohort reports of postsecondary attendance during either the third or fourth follow-up. A student may have reported attendance at more than one school. The first transcript weight was computed for students for whom at least one transcript was obtained. It is therefore possible for a student who was not a respondent in the fourth follow-up (but who was a respondent in the third follow-up) to have a nonzero value for the first transcript weight. The second transcript weight is more restrictive. It was designed to assign weights only to cases that were deemed to have complete data. Only students who responded during the fourth follow-up (and hence students for whom a complete report of postsecondary education attendance was available and for whom all requested transcripts were received) were assigned a nonzero value for the second transcript weight. For students who did not complete the fourth follow-up interview, complete transcripts may have been obtained in the 1987 transcript study, but since it was not certain that these transcripts were complete, they were given a weight of zero.

Imputation. No imputation was performed in HS&B.

5. DATA QUALITY AND COMPARABILITY

Sampling Error

Because the sample design for the HS&B cohorts involved stratification, disproportionate sampling of certain strata, and clustered probability sampling, the calculation of exact standard errors (an indication of sampling error) for survey estimates can be difficult and expensive.

Sampling error estimates for the first and second HS&B follow-ups were calculated by the method of Balanced Repeated Replication (BRR) using *BRRVAR*, a Department of Education statistical subroutine. (The BRR programs *WesVar* and *SUREG* are now available commercially.) For the base year and the third and fourth follow-ups, Taylor Series approximations were employed. More detailed discussions of the BRR and Taylor Series procedures can be found in the *High School and Beyond Third Follow-Up Sample Design Report* (Spencer et al. 1987). The Data Analysis System (DAS), included as part of the public-release file, automatically reports design-corrected Taylor Series standard errors for the tables it generates. Therefore, users of the DAS do not need to make adjustments to these estimates.

While design effects cannot be calculated for every estimate of interest to users, design effects will be similar from item to item within the same subgroup or population. Users can calculate approximate standard error estimates for items by multiplying the standard error under the simple random sample assumption by the square root of the average design effect for the population being studied.

Nonsampling Error

Nonsampling errors include coverage, nonresponse, and measurement errors.

Coverage error. Bias caused by explicit exclusion of certain groups of schools and students (e.g., special types of schools or students with disabilities or language barriers) is not addressed in HS&B technical reports. Potential coverage error in HS&B may relate to the exclusion of schools that refused to cooperate in the base-year survey. Students who refused to participate in the base-year survey were not excluded in the follow-ups. Since students were randomly selected from the sampled schools, the HS&B sample design did not entail exclusion of specified groups. (See "Sample Design," above, in section 4.)

Nonresponse error.

Unit nonresponse. HS&B base-year student-level estimates include two components of unit nonresponse bias: bias introduced by nonresponse at the school level, and bias introduced by nonresponse on the part of students attending cooperating schools. At the school level, some schools refused to participate in the base-year survey. Substitution was carried out for refusal schools within a stratum when there were two or more schools within the stratum. The bias introduced by base-year school-level refusals is of particular concern since it carried over into successive rounds of the survey. Students attending refusal schools were not sampled during the base-year and had no chance for selection into subsequent rounds of observation. To the extent that these students differed from students from cooperating schools in later waves of the study, the bias introduced by base-year school nonresponse would persist. Student nonresponse did not carry over in this way since student nonrespondents remained eligible for sampling in later waves of the study.

In general, the lack of survey data for nonrespondents prevents the estimation of unit nonresponse bias. However, during the first follow-up, School Questionnaire data were obtained from most of the base-year refusal schools, and student data were obtained from most of the base-year student nonrespondents selected for the first follow-up sample. These data provide a basis for assessing the magnitude of unit nonresponse bias in base-year estimates.

Overall, 1,120 schools were selected in the original sample, and 811 of those schools (72 percent) participated in the survey. An additional 204 schools were drawn in a replacement sample. Student refusals and absences resulted in a weighted student completion rate of 88 percent in the base-year survey. Participation was higher in most follow-up surveys. Completion rates in the first follow-up were as follows: 94 percent for seniors; 96 percent for sophomores eligible for on-campus survey administration; and 89 percent for sophomores who had left school between the base-year and first follow-up surveys (dropouts, transfer students, and early graduates). In the second follow-up, 91 percent of senior cohort members and 92 percent of sophomore cohort members completed the survey. In the third follow-up, completion rates were 88 percent for seniors and 91 percent for sophomores. Only the sophomore cohort was surveyed in the fourth follow-up; 86 percent of the sample members participated.

As results from the fourth follow-up illustrate, student nonresponse varied by demographic and educational characteristics. Males had a slightly higher

nonresponse rate than females (a difference of slightly over 3 percent). Blacks and Hispanics showed similarly high rates of nonresponse (around 20 percent), whereas nonresponse among White students was about 10 percent. Nonresponse increased as socioeconomic status decreased. Students who were in general or vocational programs during the base-year were more likely to be nonrespondents than students in academic programs. Dropouts had higher nonresponse rates than other students. Students with lower grades and lower test scores showed higher nonresponse than students with higher grades and test scores. Students who were frequently absent from school showed higher nonresponse than students absent infrequently. Students with no postsecondary education by the time of the second follow-up had higher nonresponse than students with some postsecondary education. By selected school characteristics, the highest nonresponse rates were among students from alternative public schools, schools with large enrollments, schools in urban areas, and schools in the Northeast and West.

The patterns were similar in earlier rounds of HS&B. Nonresponse analyses conducted by NORC support the following general conclusions:

- (1) The school-level bias component in HS&B estimates is small, averaging less than 2 percent for base-year and first follow-up estimates. It is probably of a similar magnitude for fourth follow-up estimates.
- (2) The student-level bias component in base-year estimates is also small, averaging about 0.5 percent for percentage estimates.
- (3) The student-level bias component in first, second, and third follow-up estimates is limited by the nonresponse rates, which were about three-fourths of the base-year rates.
- (4) The student-level bias component in the fourth follow-up estimates is limited by the nonresponse rate, which was slightly higher than the base-year rate.

The first and second conclusions together suggest that nonresponse bias is not a major contributor to error in base-year estimates. The first and third conclusions suggest that nonresponse bias is not a major contributor to error in the first, second, and third follow-up estimates either. The first and fourth conclusions suggest that the fourth follow-up nonresponse bias might be a little greater than for the previous follow-ups, but probably not by much. Each of these conclusions must be given some qualifications. The

analysis of school-level nonresponse is based on data concerning the schools, not the students attending them. The analyses of student nonresponse are based on survey data and are themselves subject to nonresponse bias. Despite these limitations, the results consistently indicate that nonresponse had a small impact on base-year and follow-up estimates.

Item nonresponse. Among students who participated in the survey, some did not complete the questionnaire or gave invalid responses to certain questions. The amount of item nonresponse varied considerably by item. For example, in the second follow-up, a very low nonresponse rate (0.1 percent) was observed for a question asking whether the respondent had attended a postsecondary institution. A much higher nonresponse rate (12.2 percent) was obtained for a question asking if the respondent had used a micro- or minicomputer in high school. Typical item nonresponse rates ranged from 3 to 4 percent.

Imputation was not used to compensate for item nonresponse in HS&B. However, an attempt was made in the fourth follow-up to reduce item nonresponse. In previous rounds, interviews were conducted by self-administered questionnaires (SAQs). Unfortunately, respondents often skipped questions incorrectly or gave unrecognizable answers. Thus, more data were missing than would have occurred through personal interviewing. In the fourth follow-up, interviewing was conducted using a CATI program. Unlike SAQs, CATI interviewing virtually eliminated missing data attributable to improperly skipped questions.

To evaluate the effectiveness of CATI interviewing, 25 items from both the third and fourth follow-up data were selected for comparison. Refusal and “don’t know” responses were considered to be missing, but legitimate skips were not. For these 25 items, the overall percentage of missing items dropped from 4.36 percent in the third follow-up to 1.88 percent in the fourth follow-up.

CATI also eliminated all multiple responses and resulted in uncodable verbatim responses for only the two income variables. In addition, more was known about the missing data in the fourth follow-up. In the third follow-up, only 7.2 percent of the missing data were classified as refusals or “don’t know” responses. In the fourth follow-up, 50.9 percent of the missing data were classified as refusals or “don’t know” responses. The fact that most of the 25 comparisons showed a “very significant” decline in missing data supports the contention that missing data were reduced in the fourth follow-up.

Measurement error. An examination of consistency between responses to the third and fourth follow-ups provides an indication of the reliability of HS&B data.

Race/ethnicity. Race/ethnicity is one characteristic of the respondents that should not change between surveys. Overall, of the 12,310 respondents who reported their race/ethnicity on both questionnaires, 93.8 percent gave the same response in both years. However, certain race/ethnicity categories (e.g., Native American) had substantially less agreement. Only 53.4 percent of the respondents who classified themselves as Native Americans during the third follow-up classified themselves as Native Americans again during the fourth follow-up.

One explanation for these discrepancies may be the change in the method of survey administration. Unlike the third follow-up, which involved self-administered questionnaires, the fourth follow-up was conducted by telephone. The questionnaires mailed during the third follow-up had the five race/ethnicity categories listed for the respondent to see. In the fourth follow-up, respondents were simply asked over the telephone, "What is your race/ethnicity?" The interviewer coded the response. It is possible that Native Americans, Hispanics, and Asian/Pacific Islanders classified themselves as Black or White (not knowing that there was a more specific category for them to choose from), hence resulting in more Blacks and Whites in the fourth follow-up results.

Marital status. In the third follow-up, respondents were asked about their marital status in the first week of February 1986. In the fourth follow-up, respondents were asked about their marital status during and since February 1986. Although both questions asked about marital status during February 1986, respondents who had a change in marital status during the last 3 weeks of February could have given a different answer in the fourth follow-up than in the third follow-up. Overall, of the 11,850 respondents who gave their marital status in both questionnaires, 95.4 percent had answers that agreed.

Unlike the race/ethnicity question, memory and timing play an important role in matching answers for marital status. In this case, the recall period for third follow-up respondents was years shorter than the recall period for respondents in the fourth follow-up. Respondents in the third follow-up, which took place in spring 1986, were asked about a recent event. Respondents in the fourth follow-up, which was conducted in spring 1992, were asked to recall their status back in February 1986. As with the race/ethnicity question, the method of administering the question differed between rounds—

namely, the question formatting had changed and the fourth follow-up used preloaded data to verify marital status.

Data Comparability

A goal of the National Longitudinal Studies Program is to allow comparative analysis of data generated in several waves of the same study as well as to enable cross-cohort comparisons with the other longitudinal studies. While the HS&B and NLS:72 studies are largely compatible, a number of variations in sample design, questionnaires, and data collection methods should be noted as a caution to data users.

Comparability within HS&B. While many data items were highly compatible across waves, the focus of the questionnaires necessarily shifted over the years in response to the changes in the cohorts' life cycle and the concerns of education policymakers. For seniors in the base-year survey and for sophomores in both the base-year and first follow-up surveys, the emphasis was on secondary schooling. In subsequent follow-ups, increasingly more items were collected dealing with postsecondary education and employment. Also, a major change in the data collection method occurred in the fourth follow-up, when CATI was introduced as the primary approach. Earlier waves used mailed questionnaires supplemented by telephone and personal interviews.

Comparability with NLS:72. The HS&B was designed to build on NLS:72 in three ways. First, the HS&B base-year survey included a 1980 cohort of high school seniors that was directly comparable to the NLS:72 cohort (1972 seniors). Replication of selected 1972 Student Questionnaire items and test items made it possible to analyze changes subsequent to 1972 and their relationship to federal education policies and programs in that period. Second, the introduction of the sophomore cohort in HS&B provided data on the many critical educational and vocational choices made between the sophomore and senior years in high school, thus permitting a fuller understanding of the secondary school experience and how it affects students. Third, HS&B expanded the NLS:72 focus by collecting data on a range of life cycle factors, such as family formation, labor force behavior, intellectual development, and social participation.

The sample design was largely similar for both HS&B and NLS:72, except that HS&B included a sophomore sample in addition to a senior sample. The questionnaires for the two studies contained a large number of identical (or similar) items dealing with secondary education and postsecondary work experience and education. The academic tests were

also highly comparable. Of the 194 test items administered to the HS&B senior cohort in the base-year, 86 percent were identical to items that had been given to NLS:72 base-year respondents. Item response theory (IRT) was used in both studies to put math, vocabulary, and reading test scores on the same scale for 1972, 1980, and 1982 seniors. With the exception of the use of CATI in the HS&B fourth follow-up, both NLS:72 and HS&B used group administration of questionnaires and tests in the earliest surveys and mailed questionnaires in the follow-ups. HS&B, however, involved more extensive efforts to supplement the mailings by telephone and personal interviews.

Comparability with NELS:88. The sample design of HS&B was also similar to that of NELS:88. In each base-year, students were selected through a two-stage stratified probability sample, with schools as the first-stage units and students within schools as the second-stage units. Because NELS:88 base-year sample members were eighth-graders in 1988, its follow-ups encompass students (both in the modal grade progression sequence and out of sequence) and dropouts. Despite similarities, however, the sample designs of the two studies differ in three major ways: (1) the NELS:88 first and second follow-ups had relatively variable, small, and unrepresentative within-school student samples, compared to the relatively uniform, large, and representative within-school student samples in the HS&B; (2) unlike the earlier study, NELS:88 did not provide a nationally representative school sample in its follow-ups; and (3) there were differences in school and subgroup sampling and oversampling strategies in the two studies. These sample differences imply differences in the respondent populations covered. (For details on NELS:88, please refer to chapter 8).

Comparability with ELS:2002. The ELS:2002 base-year and first follow-up surveys contain many data elements that are comparable to items from the HS&B. Differences in sampling rates, sample sizes, and design effects across the studies, however, affect the precision of estimation and comparability. Asian students, for example, were oversampled in ELS:2002, but not in HS&B, where their numbers were quite small. The base-year (1980) participating sample in HS&B numbered 30,030 sophomores; in contrast, 15,362 sophomores participated in the base-year of ELS:2002. Cluster sizes within schools were much larger for HS&B (on average, 30 sophomores per school) than for ELS:2002 (just over 20 sophomores per school); larger cluster sizes are better for school effects research, but carry a penalty in greater sample inefficiency. Mean design effect (a measure of sample

efficiency) is also quite variable across the studies. For example, for 10th grade, the design effect was 2.9 for HS&B, while a more favorable design effect of 2.4 was achieved for the ELS:2002 base-year. (For details on ELS:2002, please refer to chapter 9).

6. CONTACT INFORMATION

For content information on HS&B, contact:

Aurora M. D'Amico
Phone: (202) 502-7334
E-mail: aurora.damico@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Zahs, D., Pedlow, S., Morrissey, M., Marnell, P., and Nichols, B. (1995). *High School and Beyond Fourth Follow-Up Methodology Report* (NCES 95-426). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Uses of Data

Alt, M.N., and Bradby, D. (1999) *Procedures Guide for Transcript Studies* (NCES Working Paper 95-426). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S., and Baldrige, J. (1995). *National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors* (NCES Working Paper 95-05). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S., and Dowd, K. (1995). *National Education Longitudinal Study of 1988: Conducting Trend Analyses: HS&B and NELS:88 Sophomore Cohort Dropouts* (NCES Working Paper 95-07). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S., and Taylor, J. (1995). *National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data* (NCES Working Paper 95-06). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Survey Design

Frankel, M., Kohnke, L., Buonanno, D., and Tourangeau, R. (1981). *High School and Beyond Sample Design Report*. National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Jones, C., and Spencer, B.D. (1985). *High School and Beyond Second Follow-Up (1984) Sample Design Report*. National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Rock, D.A., Hilton, T.L., Pollack, J.M., Ekstrom, R.B., and Goertz, M.E. (1985). *Psychometric Analysis of the NLS-72 and the High School and Beyond Test Batteries*. National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Spencer, B.D., Sebring, P., Campbell, B., and Carroll, D. (1987). *High School and Beyond Third Follow-Up Sample Design Report* (NCES 88-402). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Tourangeau, R.E., McWilliams, H., Jones, C., Frankel, M.R., and O'Brien, F. (1983). *High School and Beyond First Follow-Up (1982) Sample Design Report*. National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Fetters, W.B., Stowe, P., and Owings, J.A. (1984). *Quality of Responses of High School Students to Questionnaire Items* (NCES 84-216). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Salvucci, S., Walter, E., Conley, V., Fink, S., and Saba, M. (1997). *Measurement Error Studies at the National Center for Education Statistics* (NCES 97-464). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Chapter 8: National Education Longitudinal Study of 1988 (NELS:88)

1. OVERVIEW

The National Education Longitudinal Study of 1988 (NELS:88) was the third major secondary education longitudinal survey sponsored by NCES. The first two surveys—the National Longitudinal Study of the High School Class of 1972 (NLS:72) and the High School and Beyond (HS&B) Longitudinal Study—examined the educational, vocational, and personal development of young people, beginning in high school. (See chapters 6 and 7 for descriptions of these studies.) The fourth high school longitudinal study, the Education Longitudinal Study of 2002 (ELS:2002), was designed to provide trend data about critical transitions experienced by students as they proceed through high school and into postsecondary education or their careers. (See chapter 8 for a description of this study.) NELS:88 provides new data about critical transitions experienced by students as they proceed from 8th grade through high school and into postsecondary education or the workforce. It expands the knowledge base of the two previous studies by surveying adolescents at an earlier age and following them into the 21st century.

The NELS:88 base-year survey included a national probability sample of 1,052 public and private 8th-grade schools, with almost 25,000 participating students across the United States. Three follow-up surveys were conducted at 2-year intervals from 1990 to 1994. In 1994 (the third follow-up), most sample members were 2 years out of high school. A fourth follow-up was conducted in 2000. In addition to surveying and testing students, NELS:88 gathered information from the parents of students, teachers, and school administrators. Furthermore, two rounds of transcript data were collected on the 8th-grade cohort. High school transcripts were collected for all participants in the school-age sample, including dropouts and early graduates. Postsecondary transcripts were collected for students who reported attending a school beyond high school.

Purpose

To provide trend data about critical transitions experienced by young people as they leave elementary school and progress through high school into postsecondary institutions or the workforce, and provide data for trend comparisons with results from NLS:72 and HS&B as well as later longitudinal studies, such as ELS: 2002.

Components

NELS:88 collected survey data from students, dropouts, parents, teachers, and school administrators. Supplementary information was gathered from high school transcripts and course offering data provided by the schools, a Base-Year Ineligible (BYI) Study, a Followback Study of Excluded Students (FSSES), a High School Effectiveness Study (HSES), and a Postsecondary Education Transcript Study. The various components are described below.

Base-Year Survey. The base-year survey was conducted during the spring school term in 1988 and included the following:

LONGITUDINAL SAMPLE SURVEY OF THE 8TH-GRADE CLASS OF 1988; BASE-YEAR SURVEY AND FOUR FOLLOW-UPS THROUGH 2000

NELS:88 collected data from:

- Students and dropouts
- School administrators
- Teachers
- Parents/guardians
- High school transcripts
- High school course offerings
- High School Effectiveness Study
- Postsecondary education transcripts

Student Questionnaire (8th-Grade Questionnaire). Students were asked to fill out a questionnaire that included items on their home background, language use, family, opinions about themselves, plans for the future, job and chores, school life, schoolwork, and activities. Students also completed a series of curriculum-based cognitive tests in four achievement areas—reading, mathematics, science, and social studies (history/government).

Parent Questionnaire. One parent of each student completed a questionnaire requesting information about both parents' background and socioeconomic characteristics, aspirations for their children, family willingness to commit resources to their children's education, the home educational support system, and other family characteristics relevant to achievement.

Teacher Questionnaire. A Teacher Questionnaire was administered to selected 8th-grade teachers responsible for instructing sampled students in two of the four test subjects—mathematics, science, English, and social studies. The questionnaire collected information in three areas: teachers' perceptions of the sampled students' classroom performances and personal characteristics; curriculum content of the areas taught; and teachers' background and activities. Two teachers were asked to respond for each student.

School Administrator Questionnaire. Completed by an official in the participating school, this questionnaire collected information about school, student, and teacher characteristics; school policies and practices; the school's grading and testing structure; school programs and facilities; parent involvement in the school; and school climate.

First Follow-up Survey. The first follow-up survey was conducted in spring 1990. It collected information from students, teachers, and school administrators, but not parents. The student sample was refreshed to be nationally representative of students enrolled in the 10th grade in spring 1990. In addition, three new components were initiated: the Dropout Questionnaire, the Base-Year Ineligible (BYI) Study, and the High School Effectiveness Study (HSES).

Students were again requested to complete a questionnaire and take cognitive tests. The Student Questionnaire collected background information and asked students about such topics as their school and home environments, participation in classes and extracurricular activities, current jobs, goals and aspirations, and opinions about themselves. Dropouts were asked similar questions in a separate *Not Currently in School Questionnaire* (or *Dropout*

Questionnaire), which also requested specific information about reason(s) for leaving school and experiences in and out of school. Dropouts were also given cognitive tests when feasible.

School administrators provided information about their high schools in the School Administrator Questionnaire, and two teachers for each student completed the Teacher Questionnaire. There were different Teacher Questionnaires for English, mathematics, science, and history. The School Administrator and Teacher Questionnaires provided information about school administration, school programs and services, curriculum and instruction, and teachers' perceptions about their students' learning.

Second Follow-up Survey. The second follow-up survey, conducted in 1992, repeated all the components of the first follow-up survey and included the Parent Questionnaire. The student sample was again refreshed to be nationally representative of students enrolled in the 12th grade in spring 1992. A new High School Transcript Study provided archival data on the academic experience of high school students. Students in high schools designated in the first follow-up for HSES were surveyed and tested again in both the main second follow-up survey and a separate HSES.

As in the previous waves, students were asked to complete a questionnaire and cognitive tests. The cognitive tests were designed to measure 12th-grade achievement and cognitive growth between 1988 and 1992 in mathematics, science, reading, and social studies (history/citizenship/geography). The questionnaire asked students about such topics as academic achievement; perceptions about their curricula and schools; family structures and environments; social relations; and aspirations, attitudes, and values relating to high school, occupations, and postsecondary education. The Student Questionnaire also contained an *Early Graduate Supplement*, which asked early graduates to document the reasons for and circumstances of their early graduation. Students who were first-time participants in NELS:88 completed a *New Student Supplement*, containing basic demographic items requested in the base year but not repeated in the second follow-up. First follow-up dropouts were resurveyed and retested. School administrators completed the School Administrator Questionnaire, and one mathematics or science teacher for each student completed the Teacher Questionnaire.

Third Follow-up Survey. The third follow-up survey, conducted in 1994, contained only the Student Questionnaire, which collected information mainly on

issues related to employment and postsecondary education. Specific content areas included academic achievement; perceptions and feelings about school and/or job; work experience and work-related training; application and enrollment in postsecondary education institutions; sexual behavior, marriage, and family; and values, leisure-time activities, volunteer activities, and voting behavior.

Fourth Follow-up Survey. The fourth follow-up survey, conducted in 2000, contained only the Student Questionnaire, which collected information mainly on issues of employment and postsecondary education. Specific content areas included academic achievement; perceptions and feelings about school and/or job; work experience and work-related training; application and enrollment in postsecondary education institutions; sexual behavior, marriage, and family; and values, leisure-time activities, volunteer activities, and voting behavior.

Supplemental Studies. The following supplemental studies were conducted during the course of the NELS:88 project:

Base-Year Ineligible (BYI) Study. The BYI Study was added to the first follow-up survey to ascertain the status of students who were excluded from the base-year survey due to a language barrier or physical or mental disability that precluded them from completing a questionnaire and cognitive tests. Any students found to be eligible at this time were included in the follow-up surveys.

Followback Study of Excluded Students (FSES). This study—a part of the second follow-up survey—was a continuation of the first follow-up BYI Study.

High School Transcript Study. This study collected high school transcripts during the second follow-up survey. Complete transcript records were collected for (1) students attending sampled schools in spring 1992; (2) dropouts (including those in alternative programs) and early graduates; and (3) sample members who were ineligible for any wave of the survey due to mental or physical disability or language barriers. The transcript data collected from schools included student-level data (e.g., number of days absent per school year, standardized test scores) and complete course-taking histories (e.g., information on credits earned; year and term a specific course was taken; and final grades). (For more information, see chapter 29, High School Transcript Studies.)

High School Effectiveness Study (HSES). To facilitate longitudinal analysis at the school level, a School Effects Augmentation was implemented in the first

follow-up survey to provide a valid probability sample of 10th-grade schools. From the pool of NELS:88 first follow-up schools, a probability subsample of 251 urban and suburban schools in the 30 largest Metropolitan Statistical Areas was selected for the HSES; 248 of these schools were HSES participants in the first follow-up. The NELS:88 national or “core” student sample in these schools was augmented to obtain a within-school representative student sample large enough to support school effects research (i.e., the effects of school policies and practices on students). These schools and students were followed up in 1992—when the majority of the students were in 12th grade—as part of both the main NELS:88 second follow-up survey and the HSES. The HSES also provided a convenient framework for a constructed-response testing experiment in 1992. The test contained four questions that required students to derive answers from their own knowledge and experience (e.g., write an explanation, draw a diagram, solve a problem). Mathematics tests were assigned to half of the schools that were willing to commit the extra time required for such testing; the other half were assigned science tests. The second follow-up HSES was also enhanced by the collection of curriculum offerings in the Course Offerings Component. (See below.)

Course Offerings Component. This component was added to the second follow-up to provide curriculum data that can serve as a baseline for studying student outcomes. The course offerings data for this component were collected from the HSES schools. These data illuminate trends when examined in conjunction with data from the transcript studies conducted as part of the 1982 HS&B and the 1987, 1990, 1994, and 1998 National Assessment of Educational Progress (NAEP).

Postsecondary Education Transcript Study. The Postsecondary Education Transcript Study was conducted as part of the fourth follow-up survey in 2000. It targeted transcripts from all U.S. postsecondary institutions attended by NELS sample members in the fourth follow-up, excluding postsecondary information collected from foreign institutions, non-degree-granting programs, and non-credit-granting institutions. The Postsecondary Education Transcript Study supplements the postsecondary education information collected in the 1994 and 2000 follow-ups by including detailed information on types of degree programs, periods of enrollment, majors or fields of study for instructional programs, specific courses taken, grades and credits attained, and credentials earned.

Periodicity

Biennial from 1988 to 1994, a fourth follow-up was conducted in 2000. The Base-Year Ineligible Study

was conducted in 1990 as part of the first follow-up; a continuation study, the Followback Study of Excluded Students, was conducted in 1992 as part of the second follow-up. The High School Effectiveness Study was conducted in the first and second follow-ups. The High School Transcript Study was implemented in the second follow-up in 1992. The Postsecondary Education Transcript Study was conducted as part of the fourth follow-up in 2000.

2. USES OF DATA

The NELS:88 project was designed to provide trend data about critical transitions experienced by students as they leave elementary school and progress through high school and into postsecondary education or the workforce. Its longitudinal design permits the examination of changes in young people's lives and the role of school in promoting growth and positive life outcomes. The project collects policy-relevant data about educational processes and outcomes, early and late predictors of dropping out, and school effects on students' access to programs and equal opportunity to learn. These data complement and strengthen state and local efforts by furnishing new information on how school policies, teacher practices, and family involvement affect student educational outcomes (e.g., academic achievement, persistence in school, and participation in postsecondary education).

NELS:88 data can be used in three ways: in cross-sectional, longitudinal, and cross-cohort analyses (by comparing NELS:88 findings with those of NLS:72, HS&B, and ELS:2002). By following young adolescents at an earlier age (beginning in 8th grade) and into the 21st century, NELS:88 expands the base of knowledge established in the NLS:72 and HS&B studies. NELS:88 first follow-up data provide a comparison point to high school sophomores 10 years earlier, as studied in HS&B. NELS:88 second follow-up data allow trend comparisons of the high school class of 1992 with the 1972 and 1980 seniors studied in NLS:72 and HS&B, respectively. The NELS:88 third follow-up allows comparisons with NLS:72 and HS&B related to postsecondary outcomes. ELS:2002 is different from NELS:88 in that the base-year sample students are 10th-graders rather than 8th-graders. With a freshened senior sample, the ELS:2002 first follow-up supports comparisons with the NELS:88 second follow-up. The ELS:2002 first follow-up academic transcript component also offers a further opportunity for a cross-cohort comparison with the high school transcript studies of NELS:88. Together, the four studies provide measures of educational attainment in the United States

and rich resources for studying the reasons for and consequences of academic success and failure.

More specifically, NELS:88 data can be used to investigate

- *transitions from elementary to secondary school*: how students are assigned to curricular programs and courses; how such assignments affect their academic performance as well as future career and postsecondary education choices;
- *academic growth over time*: family, community, school, and classroom factors that promote growth; school classroom characteristics and practices that promote learning; effects of changing family composition on academic growth;
- *features of effective schools*: school attributes associated with student academic achievement; school effects analyses;
- *the dropout process*: contextual factors associated with dropping out; movement in and out of school, including alternative high school programs;
- *the role of the school in helping the disadvantaged*: school experiences of the disadvantaged; approaches that hold the greatest potential for helping them;
- *school experiences and academic performance of language-minority students*: variation in achievement levels; bilingual education needs and experiences;
- *students' mathematics and science learning* : math and science preparation received by students; student interest in these subjects; encouragement by teachers and school to study advanced mathematics and science; and
- *transitions from high school to college and postsecondary access/choice*: planning and application behaviors of the high school class of 1992; subsequent enrollment in postsecondary institutions.

3. KEY CONCEPTS

Some of the key terms related to NELS:88 are defined below.

Cognitive Test Battery. Cognitive tests measuring student achievement in mathematics, reading, science, and social studies (history/citizenship/geography) were administered in the base year, first follow-up, and second follow-up. The contents was as follows: (1) reading (21 items, 21 minutes); (2) mathematics (40 items, 30 minutes); (3) science (25 items, 20 minutes); and (4) social studies (30 items, 14 minutes—the base-year test included history and government items; the first and second follow-up tests included history, citizenship, and geography items).

Socioeconomic Status (SES). A composite variable constructed from five questions in the Parent Questionnaire: father's education level, mother's education level, father's occupation, mother's occupation, and family income. When all parent variables were missing, student data were used to compute the SES, substituting household items (e.g., dictionary, computer, more than 50 books, washing machine, calculator) for the family income variable. There are separate SES variables derived from parent data in the base year and the second follow-up. The database also included variables for SES quartiles.

Dropout. Used both to describe an event (leaving school before graduating) and a status (an individual who was not in school and not a graduate at a defined point in time). The NELS:88 "cohort dropout rate" is based on a measurement of the enrollment status of 1988 8th-graders 2 and 4 years later (in spring 1990 and spring 1992) and of 1990 sophomores 2 years later (in spring 1992). For a given point in time, a respondent is considered to be a dropout if he or she had not graduated from high school or attained an equivalency certificate and had not attended high school for 20 consecutive days (not counting excused absences). Transferring to another school is not regarded as a dropout event, nor is delayed graduation if a student was continuously enrolled but took an additional year to complete high school. A person who dropped out of school may have returned later and graduated. This person would be considered a "dropout" at the time he or she initially left school and a "stopout" at the time he or she returned to school.

4. SURVEY DESIGN

Target Population

Students enrolled in the 8th grade in "regular" public and private schools located in the 50 states and the District of Columbia in the spring 1988 school term. The sample was freshened in both the first and second follow-ups to provide valid probability samples that would be nationally representative of 10th-graders in spring 1990 and 12th-graders in spring 1992. The NELS:88 project excludes the following types of schools: Bureau of Indian Education (BIE)¹ schools, special education schools for the handicapped, area vocational schools that do not enroll students directly, and schools for dependents of U.S. personnel overseas. The following students are also excluded: mentally handicapped students and students not proficient in English, for whom the NELS:88 tests would be unsuitable; and students having physical or emotional problems that would make participation in the survey unwise or unduly difficult. However, the Base-Year Ineligible Study (in the first follow-up) and the Followback Study of Excluded Students (in the second follow-up) sampled excluded students and added those no longer considered ineligible to the freshened sample of the first and second follow-ups, respectively.

Sample Design

NELS:88 was designed to follow a nationally representative longitudinal component of students who were in the 8th grade in spring 1988. It also provides a nationally representative sample of schools offering 8th grade in 1988. In addition, by freshening the student sample in the first and second follow-ups, NELS:88 provides nationally representative populations of 10th-graders in 1990 and 12th-graders in 1992. To meet the needs for cross-sectional, longitudinal, and cross-cohort analyses, NELS:88 involved complex research designs, including both longitudinal and cross-sectional sample designs.

Base-Year Survey. In the base year, students were selected using a two-stage stratified probability design, with schools as the first-stage units and students within schools as the second-stage units. From a national frame of about 39,000 schools with 8th grades, a pool of 1,030 schools was selected through stratified sampling with probability of selection proportional to their estimated 8th-grade enrollment; private schools were oversampled to ensure adequate representation. A pool of 1,030 replacement schools was selected by the same method to be used as substitutions for ineligible or

¹ These were referred to as Bureau of Indian Affairs (BIA) funded schools.

refusal schools in the initial pool. A total of 1,060 schools cooperated in the base year; of these, 1,060 schools (815 public and 237 private) contributed usable student data. The sampling frame for NELS:88 was the school database compiled by Quality Education Data, Inc., of Denver, Colorado, supplemented by racial/ethnic data obtained from the U.S. Office for Civil Rights and school district personnel.

Student sampling produced a random selection of 26,440 8th-graders in 1988; 24,600 participated in the base-year survey. Hispanic and Asian/Pacific Islander students were oversampled. Within each school, approximately 26 students were randomly selected (typically, 24 regularly sampled students and 2 oversampled Hispanic or Asian/Pacific Islander students). In schools with fewer than 24 8th-graders, all eligible students were selected. Potential sample members were considered ineligible and excluded from the survey if disabilities or language barriers were seen as obstacles to successful completion of the survey. The eligibility status of excluded members was reassessed in the first and second follow-ups. (See below.)

First Follow-up Survey. There were three basic objectives for the first follow-up sample design. First, the sample was to include approximately 21,500 students who were in the 8th-grade sample in 1988 (including base-year nonrespondents), distributed across 1,500 schools. Second, the sample was to constitute a valid probability sample of all students enrolled in the 10th grade in spring 1990. This entailed “freshening” the sample with students who were 10th-graders in 1990 but who were not in the 8th grade in spring 1988 or who were out of the country at the time of base-year sampling. The freshening procedure added 1,230 10th-graders; 1,040 of the students in this new group were found to be eligible and were retained after final subsampling for the first follow-up survey. Third, the first follow-up was to include a sample of students who had been deemed ineligible for base-year data collection due to physical, mental, or linguistic barriers to participation. The Base-Year Ineligible Study reassessed the eligibility of these students so that those able to take part in the survey could be added to the first follow-up student sample. Demographic and school enrollment information was also collected for all students excluded in the base year, regardless of their eligibility status for the first follow-up.

While schools covered in the NELS:88 base-year survey were representative of the national population of schools offering the 8th grade, the schools in the first follow-up were not representative of the national population of high schools offering the 10th grade. By 1990, the 1988 8th-graders had dispersed to many high

schools, which did not constitute a national probability sample of high schools. To compensate for this limitation, the High School Effectiveness Study (HSES), which was designed to sustain analyses of school effectiveness issues, was conducted in conjunction with the first follow-up. From the pool of participating first follow-up schools, a probability subsample of 251 urban and suburban schools in the 30 largest Metropolitan Statistical Areas were designated as HSES schools. The NELS:88 core student sample was augmented to obtain a within-school representative student sample large enough to support school effects research. The student sample was increased in HSES schools by an average of 15 students to obtain within-school student cluster sizes of approximately 30 students.

Second Follow-up Survey. The second follow-up sample included all students and dropouts selected in the first follow-up. From within the schools attended by the sample members, 1,500 12th-grade schools were selected as sampled schools. Of these, the full complement of component activities occurred in 1,370 schools. For students attending schools other than these 1,370 schools, only the Student and Parent Questionnaires were administered. As in the first follow-up, the student sample was augmented through freshening to provide a representative sample of students enrolled in the 12th grade in spring 1992. Freshening added into the sample 243 eligible 12th-graders who were not in either the base-year or first follow-up sampling frames. Schools and students designated for the HSES in the first follow-up were followed up again—as part of both the main second follow-up survey and a separate HSES. The Followback Study of Excluded Students was a continuation of the first follow-up Base-Year Ineligible Study. In addition, two new components—the High School Transcript Study and the Course Offerings Component—were added to the second follow-up.

Third Follow-up Survey. The third follow-up student sample was created by dividing the second follow-up sample into 18 groups based on students’ response history, dropout status, eligibility status, school sector type, race, test scores, SES, and freshened status. Each sampling group was assigned an overall selection probability. Cases within a group were selected such that the overall group probability was met, but the probability of selection within the group was proportional to each sample member’s second follow-up design weight. Assigning selection probabilities in this way reduced the variability of the third follow-up raw weights and consequently increased the efficiency of the resulting sample from 40.1 to 44.0 percent.

Fourth Follow-up Survey. The fourth follow-up student sample was the same as the third follow-up student sample. Data collection for the NELS:88 fourth follow-up survey ended in September 2000, providing a final respondent population of approximately 12,100 individuals.

The Postsecondary Education Transcript Study, conducted as part of the fourth follow-up in 2000, followed those who reported having attended at least one postsecondary institution according to either the third follow-up survey in 1994 or the fourth follow-up survey in 2000. A total of approximately 9,600 fourth follow-up survey respondents (79 percent of the overall respondent population) reported postsecondary experience since high school. Approximately 21 percent of the NELS:88 respondent population did not participate in postsecondary education.

Within this sample of students, the transcript data collection further targeted students who attended only postsecondary institutions identified in the Integrated Postsecondary Education Data System (IPEDS) institutional data file, thus excluding postsecondary information collected from foreign institutions, non-degree-granting programs, and non-credit-granting institutions. Transcripts were requested from a total of 3,200 postsecondary institutions.

Data Collection and Processing

NELS:88 compiled data from five primary sources: students, parents, school administrators, teachers, and high school administrative records (transcripts, course offerings, and course enrollments). Data collection efforts for the base year through third follow-up extended from spring 1988 through summer 1994. Self-administered questionnaires, cognitive tests, and telephone or personal interviews were used to collect the data. The follow-up surveys involved extensive efforts to locate and collect data from sample members who were school dropouts, school transfers, or otherwise mobile individuals. Coding and editing conventions adhered as closely as possible to the procedures and standards previously established for NLS:72 and HS&B. The contractor National Opinion Research Center (NORC) at the University of Chicago was the prime contractor for the NELS:88 project from the base year through the third follow-up, but Research Triangle Institute conducted the fourth follow-up.

Reference dates. In the base-year survey, most questions referred to the student's experience up to the time of the survey administration in spring 1988. In the follow-ups, most questions referred to experiences that occurred between the previous survey and the current survey. For example, the second follow-up largely

covered the period between 1990 (when the first follow-up was conducted) and 1992 (when the second follow-up was conducted).

Data collection. Prior to each survey, it was necessary to secure a commitment to participate in the study from the administrator of each sampled school. For public schools, the process began by contacting the Council of Chief State School Officers and the officer in each state. Once approval was gained at the state level, contact was made with district superintendents and then with school principals. For private schools, the National Catholic Educational Association and the National Association of Independent Schools were contacted for endorsement of the project, followed by contact of the school principals. The principal of each cooperating school designated a School Coordinator to serve as a liaison between contractor staff and selected respondents—students, parents, teachers, and the school administrator. The School Coordinator (most often a guidance counselor or senior teacher) handled all requests for data and materials, as well as all logistical arrangements for student-level data collection on the school premises. Coordinators were asked to identify students whose physical or learning disabilities or linguistic deficiencies would preclude participation in the survey and to classify all eligible students as White, Black, Hispanic, Asian/Pacific Islander, or “other” race.

For the base-year through second follow-up surveys, Student Questionnaires and test batteries were primarily administered in group sessions at the schools on a scheduled Survey Day. The sessions were monitored by contractor field staff, who also checked the questionnaires for missing data and attempted data retrieval while the students were in the classroom. Makeup sessions were scheduled for students who were unable to attend the first session. In the first and second follow-ups, off-campus sessions were used for dropouts and for sample members who were not enrolled in a first follow-up school on Survey Day. The School Administrator, Teacher, and Parent Questionnaires were self-administered. Contractor field staff followed up by telephone with individuals who had not returned their questionnaires by mail within a reasonable amount of time.

The first follow-up data collection required intensive tracing efforts to locate base-year sample members who, by 1990, were no longer in their 8th-grade schools but had dispersed to many high schools. Also, in order to derive a more precise dropout rate for the 1988 8th-grade cohort, a second data collection was undertaken 1 year later, in spring 1991. At this time, an attempt was made to administer questionnaires—by telephone or in person—to sample members who had missed data

collection at their school or who were no longer enrolled in school. The first follow-up also included the Base-Year Ineligible (BYI) Study, which surveyed a sample of students considered ineligible in the base year due to linguistic, mental, or physical deficiencies. The BYI Study sought to determine if eligibility status had changed for the excluded students so that newly eligible students could be added to the longitudinal sample. If an excluded student was now eligible, an abbreviated Student Questionnaire or a Dropout Questionnaire was administered, as appropriate. For those students who were still ineligible, their school enrollment status was ascertained and basic information about their sociodemographic characteristics was recorded.

Tracing efforts continued in the second and third follow-ups. In the second follow-up (conducted in 1992), previously excluded students were surveyed through the Followback Study of Excluded Students. The second follow-up also collected transcript, course offerings, and course enrollments from the high schools; reminder postcards were sent to principals who did not respond within a reasonable period. Data collection for the High School Effectiveness Study (HSES) was conducted concurrently with the collection for the second follow-up. Because of the overlap in school and student samples, survey instruments and procedures for the HSES were almost identical to those used in the NELS:88 second follow-up survey.

By 1994, when the third follow-up was conducted, most sample members had graduated from high school and it was no longer feasible to use group sessions to administer Student Questionnaires. Instead, the dominant form of data collection was one-on-one administration through computer-assisted telephone interviewing (CATI). In-person interviews were used for sample members who required intensive in-person locating or refusal conversion. Only the Student Questionnaire was administered in the third follow-up.

By 2000, when the fourth follow-up was conducted, most sample members who attended college and technical schools had completed their postsecondary education. Data collection for the fourth follow-up survey was conducted almost exclusively with computer-assisted interviewing, primarily by telephone (i.e., using CATI). However, in-person field interviews were also completed with this technology. Field interviewers used the same computer-assisted interview and online coding software as the study's telephone interviewers, but on a laptop computer-based platform (i.e., computer-assisted personal interviewing, or CAPI). Thus, all of the entry of interview data was

accomplished by the NELS:88 fourth follow-up CATI-CAPI system.

High school transcripts were collected as part of the second follow-up. The groundwork for the collection of high school transcripts was laid in the spring and fall of 1991, during pre-data collection activities for the second follow-up. Principals were asked to provide any materials—such as course catalogs, student manuals or handbooks, course lists, and registration forms—that would aid transcript course coding. In August 1992, transcript survey materials were mailed to the principals of the NELS:88 and non-NELS:88 schools attended or most recently attended by sample members eligible for the survey. Two weeks after survey materials were mailed, nonresponding principals were prompted for the return of transcripts with a postcard reminder. Principals who did not return transcripts within 3 weeks of the postcard prompt were prompted over the telephone. Telephone prompting of nonresponding principals continued from October 1992 through February 1993. Field visits to schools requesting assistance in the preparation of transcripts were conducted in February and March.

The Postsecondary Education Transcript Study was carried out at the conclusion of CATI and CAPI data collection for the fourth follow-up survey. Data collection began in September 2000, and over the next 5 months project staff requested transcripts from postsecondary institutions that NELS:88 fourth follow-up respondents reported attending during either the NELS:88 third follow-up or NELS:88 fourth follow-up studies. Requests for transcripts were sent to the registrars or other contacts at the schools. Telephone follow-up with nonresponding institutions took place 2 weeks after transmission of the package. Data collection procedures were designed to follow, where possible, each institution's typical procedures for producing and distributing student transcripts. Returned transcripts and related school catalogs and bulletins were inventoried, transcript identification numbers affixed to each, and unique identifying information removed.

Processing. Data processing activities were quite similar for the base-year survey and the first and second follow-ups. An initial check of student documents for missing data was performed on-site by contractor staff so that data could be retrieved from the students before they left the classroom. Special attention was paid to a list of "critical items." Once the questionnaires and tests were received at the contractor, they were again reviewed for completeness, and a final disposition code was assigned to the case indicating which documents had been completed by the

sample member. Postsecondary institutions reported by the student were coded using the standard IPEDS codes. Data entry for both Student Questionnaires and cognitive tests was performed through optical scanning. New Student Supplements and Dropout Questionnaires were converted to machine-readable form using key-to-disk methods. All cognitive tests were photographed onto microfilm for archival storage. In the third follow-up, a CATI system captured the data at the time of the interview. The system evaluated the responses to completed questions and used the results to route the interviewer to the next appropriate question. The CATI program also applied the customary edits, described below under "Editing." At the conclusion of an interview, the completed case was deposited in the database ready for analysis. There was minimal post-data entry cleaning because the interviewing module itself conducted the majority of necessary edit checking and conversion functions.

Verbatim responses were collected in the third follow-up for a number of items, including occupation and major field of study. When respondents indicated their occupation, the CATI interviewers recorded the verbatim response. The system checked the response using a keyword search to match it to a subset of standard industry and occupation codes, and then presented the interviewer with a set of choices based on the keyword matches. The interviewer chose the option which most closely matched the information provided by the respondent, probing for additional information when necessary. Quality control was ensured by a reading and recoding, if necessary, of the verbatim responses by professional readers.

In the fourth follow-up, data were collected and edited almost exclusively with computer-assisted interviewing, primarily by telephone (i.e., using CATI).

For the High School Transcript Study, student- and course-level data were abstracted from transcripts. Transcript courses were coded using the course catalog for the school or district, in accordance with the Classification System of Secondary Courses, updated for the 1990 NAEP High School Transcript Study. When a school or district catalog was unavailable, courses were coded by title alone.

Information from the postsecondary education transcripts, including terms of attendance, fields of study, specific courses taken, and grades and credits earned, was coded and processed using a transcript control system developed specifically for this purpose. Specially trained research personnel then coded and tabulated these academic documents.

Editing. In the base-year through second follow-up surveys, detection of out-of-range codes was completed during scanning or data entry for all closed-ended questions. Machine editing was used to (1) resolve inconsistencies between filter and dependent questions; (2) supply appropriate missing data codes for questions left blank (e.g., legitimate skip, refusal); (3) detect illegal codes and convert them to missing data codes; and (4) investigate inconsistencies or contradictions. Frequencies and cross-tabulations for each variable were inspected before and after these steps to verify the accuracy and appropriateness of the machine editing. Items with unusually high nonresponse or multiple responses were further checked by verifying the responses on the questionnaire. A final editing step involved recoding Student Questionnaire responses for some items to the codes for the same items in earlier NELS:88 waves or in HS&B. Once this was done, codes that differed in the Dropout Questionnaire were recoded to coincide with the codes used for Student Questionnaire responses.

In the third and fourth follow-ups, machine editing was replaced by the interactive edit capabilities of the CATI system, which tested responses for valid ranges, data field size, data type (numeric or text), and consistency with other answers or data from previous rounds. If the system detected an inconsistency because of an interviewer's incorrect entry, or if the respondent simply realized that he or she had made a reporting error earlier in the interview, the interviewer could go back and change the earlier response. As the new response was entered, all of the edit checks performed at the first response were again performed. The system then worked its way forward through the questionnaire using the new value in all skip instructions, consistency checks, and the like until it reached the first unanswered question, and control was then returned to the interviewer. When problems were encountered, the system could suggest prompts for the interviewer to use to elicit a better or more complete answer.

Estimation Methods

Sample weighting is required so that NELS:88 data are representative of the full population. Imputation for missing nonresponses, however, has not yet been systematically provided for data analysis.

Weighting. Weighting is used in NELS:88 data analysis to accomplish a number of objectives, including (1) expanding counts from sample data to full population levels; (2) adjusting for differential selection probabilities (e.g., the oversampling of Asian and Hispanic students); (3) adjusting for differential response rates; and (4) improving representativeness by using auxiliary information. Multiple "final" (or

nonresponse-adjusted) weights have been provided for analyzing the different populations that NELS:88 data represent (i.e., base-year schools; 8th-graders in 1988 and 2, 4, 6, and 12 years later; 1990 sophomores; 1992 seniors; and 2000 college graduates). Weights should be used together with the appropriate flag in order to analyze the sample for a particular targeted population.

Weights have not been constructed for all possible analytic purposes. In cases where no specific weight is available, existing weights may provide reasonable approximations. For instance, base-year parent and cognitive test completion rates were so high relative to Student Questionnaire completion that the student weight can be used for them with minimal bias.

NELS:88 weights were calculated in two steps: (1) unadjusted weights were calculated as the inverse of the probabilities of selection, taking into account all stages of the sample selection process; and (2) these initial weights were adjusted to compensate for nonresponse, typically carried out separately within multiple weighting cells. For detailed discussions of the calculation of weights for each wave, users are referred to the methodology reports for the individual surveys.

Scaling (Item Response Theory). Item Response Theory (IRT) was used to calibrate item parameters for all cognitive test items administered to students in NELS:88 tests. The tests conducted in each NELS:88 survey generated achievement measures in standardized scores.

Imputation. NELS:88 surveys have not involved large-scale imputation of missing data. Only a few variables have been imputed: student's sex, race/ethnicity, and school enrollment status. For example, when sex was missing in the data file, the information was looked for in earlier school rosters. If it was still unavailable after this review, sex was assumed from the sample member's name (if unambiguous). As a final resort, sex was randomly assigned.

5. DATA QUALITY AND COMPARABILITY

A number of studies have been conducted to address data quality issues relating to the NELS:88 project. During the course of data collection and processing, systematic efforts were made to monitor, assess, and maximize data quality. Subsequently, studies were conducted to evaluate the data quality in NELS:88 in comparison with that in earlier longitudinal surveys.

Sampling Error

Because the NELS:88 sample design involved stratification, disproportionate sampling of certain strata, and clustered (i.e., multistage) probability sampling, the calculation of exact standard errors (an indication of sampling error) for survey estimates can be difficult and expensive. For NELS:88, the Taylor series procedure has typically been used to calculate the standard errors.

Standard errors and design effects for about 30 key variables in each NELS:88 wave from the base year through the fourth follow-up were calculated using SUDAAN software. These can be used to approximate the standard errors if users do not have access to specialized software.

Design effects. A comparative study of design effects across NELS:88 waves and between NELS:88 and HS&B was done. When comparing NELS:88 base-year Student Questionnaire data to the results from HS&B—the 30 variables from the NELS:88 Student Questionnaire were selected to overlap as much as possible with those variables examined in HS&B—the design effects indicate that the NELS:88 sample was slightly more efficient than the HS&B sample. The smaller design effects in the NELS:88 base year may reflect its smaller cluster size (24 students plus, on average, two oversampled Hispanics and Asian from each NELS:88 school vs. the 36 sophomore and 36 senior selections from each HS&B school). The mean design effect for base-year students is 2.54.

In the comparative study of design effects across NELS:88 waves, the design effects in the subsequent follow-up studies were somewhat higher than those in the base year, a result of the subsampling procedures used in the follow-ups. The mean design effects for students and dropouts are 3.90 for the first follow-up, 3.70 for the second follow-up, 2.90 for the third follow-up, and 3.90 for the fourth follow-up. See the *NELS:88 Base Year Through Second Follow-up Final Methodology Report* (Ingels et al. 1998) and the *User's Manual: NELS:88 Base-Year to Fourth Follow-up: Student Component Data File* (Curtin et al. 2002).

Nonsampling Error

Coverage error. Exclusion and undercoverage of certain groups of schools and students in NELS:88 generated coverage error. In the base-year survey, for example, students who had linguistic, mental, or physical obstacles were excluded from the study.

Consequently, the national populations for such student groups were not fully covered by the sample.

To correct this coverage bias, the Base-Year Ineligible (BYI) Study collected eligibility information for 93.9 percent of the sample members excluded in the base-year survey. For those who were reclassified as eligible in the BYI Study, Student or Dropout Questionnaires were administered in person or over the telephone during the first follow-up. Cognitive tests were also administered to a small percentage of these students. For students who remained ineligible, school enrollment status and other key characteristics were obtained. The BYI Study permitted an evaluation of coverage bias in NELS:88 and a means of reducing undercoverage by identifying newly eligible students who could then be added into the sample to ensure cross-sectional representativeness. This effort also provided a basis for making corrected dropout estimates, taking into account both 1988-eligible and 1988-ineligible 8th-graders 2 years later. For further detail on the BYI Study, see *Sample Exclusion in NELS:88: Characteristics of Base Year Ineligible Students; Changes in Eligibility Status After Four Years* (Ingels 1996).

Nonresponse error. Both unit nonresponse (nonparticipation in the survey by a sample member) and item nonresponse (missing value for a given questionnaire/test item) have been evaluated in

NELS:88 data.

Unit nonresponse. In the NELS:88 base-year survey, the initial school response rate was 69 percent. This low rate prompted a follow-up survey to collect basic characteristics from a sample of the nonparticipating schools. These data were then compared to the same characteristics among the participating schools to assess the possible impact of response bias on the survey estimates. The school-level nonresponse bias was found to be small to the extent that schools could be characterized by size, control, organizational structure, student composition, and other factors. Bias at the school level was not assessed for the follow-up surveys because (1) sampling for the first and second follow-ups was student-driven (i.e., the schools were identified by following student sample members) and the third and fourth follow-ups did not involve schools; and (2) school cooperation rates were very high (up to 99 percent). Even if a school refused to cooperate, individual students were pursued outside of school (although school context data were not collected). The student response rates are shown in table 5 below.

Student-level nonresponse analysis was conducted with a focus on *panel nonresponse* since a priority of

Table 5. Unit-level and overall weighted response rates for selected NELS:88 student populations, by data collection wave

Population	Unit-level weighted response rate					
	Base-year school level	Base-year student level	1st follow-up	2nd follow-up	3rd follow-up	4th follow-up
Interviewed students	69.7 ¹	93.4	91.1	91.0	90.9	82.1
Tested students	69.7 ¹	96.5	94.1	76.6	†	†
Dropouts	69.7 ¹	†	91.0	88.0	†	†
Tested dropouts	69.7 ¹	†	48.6	41.7	†	†
Population	Overall weighted response rate					
	Base-year school level	Base-year student level	1st follow-up	2nd follow-up	3rd follow-up	4th follow-up
Interviewed students	69.7 ¹	65.1	63.5	63.4	63.4	57.2
Tested students	69.7 ¹	67.3	65.6	53.4	†	†
Dropouts	69.7 ¹	†	63.4	61.3	†	†
Tested dropouts	69.7 ¹	†	33.9	29.1	†	†

† Not applicable.

¹Unweighted response rate.

SOURCE: Curtin, T.R., Ingels, S.J., Wu, S., and Heuer, R. (2002). *User's Manual: NELS:88 Base-Year to Fourth Follow-up: Student Component Data File* (NCES 2002-323). National Center for Education Statistics, U.S. Department of Education, Washington, DC. Spencer, B.D., Frankel, M.R., Ingel, S.J., Rasinski, K.A., and Tourangeau, R. (1990). *NELS:88 Base-Year Sample Design Report* (NCES 90-463). National Center for Education Statistics, U.S. Department of Education, Washington, DC.

the NELS:88 project is to provide a basis for longitudinal analysis. Nonresponse was examined for the 8th-grade and 10th-grade cohorts. Any member of the 8th-grade cohort who did not complete a survey in three rounds (base year, first follow-up, and second follow-up) and any member in the 10th-grade cohort who did not complete a survey in the second and third rounds (first and second follow-ups) was considered a panel nonrespondent for that cohort. Panel nonresponse to cognitive tests in the two cohorts was defined the same way. The nonresponse rate was defined as the proportion of the selected students (excluding deceased students) who were nonrespondents in any round in which data were collected.

Nonresponse rates for both cohorts were calculated by school- and student-level variables that were assumed to be stable across survey waves (e.g., sex and race). These variables allowed comparisons between participants and nonparticipants even though the data for the latter were missing in some rounds. Estimates were made with both weighted and unweighted data. The weight used was the second follow-up raw panel weight (not available in the public-release dataset). About 18 percent of the 8th-grade cohort and 10 percent of the 10th-grade cohort were survey nonrespondents at one or more points in time. Approximately 43 percent of the 8th-grade cohort and 35 percent of the 10th-grade cohort did not complete one or more cognitive tests in their rounds of testing.

Nonresponse bias was calculated as the difference in the estimates between the respondents and all selected students. On the whole, the analysis revealed only small discrepancies between the two cohorts. Bias estimates were higher, however, for the 8th-grade cohort than for the 10th-grade cohort because of the 8th-grade cohort's more stringent definition of participation. The discrepancies between cognitive test completers and noncompleters were larger than between survey participants and nonparticipants; this pattern held for both cohorts. In brief, the magnitude of the bias was generally small—few percentage estimates were off by as much as 2 percent in the 8th-grade cohort and 1 percent in the 10th-grade cohort. Such bias reflects the raw weight. The nonresponse-adjusted weight should correct for differences by race and sex to produce correct population estimates for each subgroup.

Further analysis was done using several other student and school variables. The results showed rather similar patterns of bias. When compared with estimates from HS&B, the student nonresponse bias estimates in NELS:88 were consistently lower. However, the two

studies seem to share certain common patterns of nonresponse. For example, both studies generated comparatively higher nonresponse rates among students enrolled in schools in the West, Black students, students in vocational or technical programs, students in the lowest test quartile, and dropouts.

Item nonresponse. Item nonresponse was examined in base-year through second follow-up data obtained from surveys of students, parents, and teachers. Differences emerged among student subgroups in the level of nonresponse to a wide range of items—from language background, family composition, and parents' education to perception of school safety. Nonresponse was often two to five times as great for one subgroup as for the other subgroups. High item nonresponse rates were associated with such attributes as not living with parents, having low SES, being male, having poor reading skills, and being enrolled in a public school. Compared with parent nonresponse to items about college choice and occupational expectations, student nonresponse rates were generally lower. For items about student's language proficiency, classroom practices, and student's high school track, students had consistently lower nonresponse rates than their teachers did. See the *NELS:88 Survey Item Evaluation Report* (McLaughlin, Cohen, and Lee 1997) for further detail.

Measurement error. NCES has conducted studies to evaluate measurement error in (1) student data (compared to parent and teacher data); and (2) student cognitive test data.

Parent-student convergence and teacher-student convergence. A study of measurement error in data from the base-year through second follow-up surveys focused on the convergence of responses by parents and students and by teachers and students. (See the *NELS:88 Survey Item Evaluation Report* [McLaughlin, Cohen, and Lee 1997].) Response convergence (or discrepancy) across respondent groups can be interpreted as an indication of measurement reliability, validity, and communality, although the data are often not sufficient to determine which response is more accurate.

The student and parent components of this study covered such variables as number of siblings, the student's work experience, language background, parents' education, parent-student discussion of issues, perceptions about school, and college and occupation expectations. Parent-student convergence varied from very high to very low, depending on the item. For example, convergence was high for number of siblings, regardless of student-level characteristics such as SES, sex, reading scores, public versus private school

enrollment, and whether or not living with parents. In contrast, parent-student convergence was low for items related to the student's work experience; there was also more variation across student subgroups for these items. In general, convergence tended to be high for objective items, for items worded similarly, and for nonsensitive items.

Teacher-student convergence was examined through variables about student's English proficiency, classroom practices, and student's high school track. Again, convergence was found to vary considerably across data items and student subgroups. Convergence was high for student's native language but low for student's English proficiency. Across student subgroups, there was a greater range in correlations for English proficiency than for native language. Teachers and students differed quite dramatically on items about classroom practices.

Cognitive test data. In-depth studies of measurement error issues related to cognitive tests administered in the base-year through second follow-up surveys are also available. See the *Psychometric Report for the NELS:88 Base Year Test Battery* (Rock and Pollack 1991) and the *Psychometric Report for the NELS:88 Base Year Through Second Follow-up* (Rock and Pollack 1995).

The first study (Rock and Pollack 1991) addressed issues related to test speediness (the limited testing time in relation to the outcome), reliability, item statistics, performance by racial/ethnic and gender groups, and IRT parameters for the battery. The results indicate that the test battery either met or exceeded all of its psychometric objectives. Specifically, the study reported: (1) while the allotted testing time was only 1½ hours, quite acceptable reliability was obtained for the tests on reading comprehension, mathematics, history/citizenship/ geography, and, to a somewhat lesser extent, science; (2) the internal consistency reliability was sufficiently high to justify the use of IRT scoring and, thus, provide the framework for constructing 10th- and 12th-grade forms that would be adaptive to the ability levels of the students; (3) there was no consistent evidence of differential item functioning (item bias) for gender or racial/ethnic groups; (4) factor analysis results supported the discriminant validity of the four tested content areas; convergent validity was also indicated by salient loadings of testlets composed of "marker items" on their hypothesized factors; and (5) in addition to providing the usual normative scores in all four tested areas, behaviorally anchored proficiency scores were provided in both the reading and math areas.

The second study (Rock and Pollack 1995) focused on issues relating to the measurement of gain scores. Special procedures were designed into the test battery design and administration to minimize the floor and ceiling effects that typically distort gain scores. The battery used a two-stage multilevel procedure that attempted to tailor the difficulty of the test items to the performance level of a particular student. Thus, students who performed very well on their 8th-grade mathematics test received a relatively more difficult form in 10th grade than students who had not performed well on their 8th-grade test. There were three forms of varying difficulty in mathematics and two in reading in both grades 10 and 12. Since 10th- and 12th-graders were taking forms that were more appropriate for their level of ability and achievement, measurement accuracy was enhanced and floor and ceiling effects could be minimized. The remaining two content areas—science and history/citizenship/geography—were only designed to be grade-level adaptive (i.e., a different form for each grade but not multiple forms varying in difficulty within grade).

To maximize the gain from using an adaptive procedure, special vertical scaling procedures were used that allow for Bayesian priors on subpopulations for both item parameters and scale scores. In comparing more traditional non-Bayesian approaches to scaling longitudinal measures with the Bayesian approach, it was found that the multilevel approach did increase the accuracy of the measurement. Furthermore, when used in combination with the Bayesian item parameter estimation, the multilevel approach reduced floor and ceiling effects when compared to the more traditional IRT approaches.

Data Comparability

NELS:88 is designed to facilitate both longitudinal and trend analyses. Longitudinal analysis calls for data compatibility across survey waves whereas trend analysis requires data compatibility with other longitudinal surveys. Data compatibility issues may relate to survey instruments, sample design, and data collection methods.

Comparability within NELS:88 across survey waves.

A large number of variables are common across survey waves. (See the *NELS:88 Second Follow-up Student Component Data File User's Manual* [Ingels et al. 1994] for a listing of common Student Questionnaire variables in the base year, first follow-up, and second follow-up.) However, compatibility of NELS:88 data across waves can still be an issue because of subtle differences in question wording, sample differences (e.g., with or without dropouts and freshening students, sample attrition, nonresponse), and data collection

methods (e.g., on-campus group session, off-campus individual survey, telephone interview).

One NCES study compared 112 pairs of variables repeated from the base year to the first and second follow-up surveys. (See the *NELS:88 Survey Item Evaluation Report* [McLaughlin, Cohen, and Lee 1997].) These variables cover student family, attitudes, education plans, and perceptions about schools. The results suggest that the interpretations of NELS:88 items depend on the age level at which they were administered. Data convergence tended to be higher for pairs of first and second follow-up measures than for pairs of base-year and second follow-up measures. Some measures were more stable than others. Students responded nearly identically to the base-year and second follow-up questions about whether English was their native language. Their responses across survey waves were also fairly stable as to whether their curriculum was intended to prepare them for college, whether they planned to go to college, and their religiosity. It should be noted that cross-wave discrepancies may reflect a change in actual student behavior rather than a change in response for a status quo situation.

Comparability within NELS:88 across respondent groups. While different questionnaires were used to collect data from different respondent groups (students, parents, teachers, school administrators), there are overlapping items among these instruments. One study examined the extent to which the identical or similar items in different questionnaires generated compatible information. It found considerable discrepancies between students and parents, and even greater discrepancies between students and teachers, in their responses to selected groups of overlapping variables. (See “Measurement error” above.)

Comparability with NLS:72, HS&B, and ELS:2002. NELS:88 surveys contain many items that are also covered in NLS:72, HS&B, and ELS:2002—a feature that enables trend analyses of various designs. (See the *NELS:88 Second Follow-up Student Component Data File User’s Manual* [Ingels et al. 1994] for a cross-walk of common variables and a discussion of trend analyses.) To examine data compatibility across the four studies, one should consider their sample designs and data contents, including questionnaires, cognitive tests, and transcript records.

Sample designs for the four studies are similar. In each base year, students were selected through a two-stage stratified probability sample, with schools as the first-stage units and students within schools as the second-stage units. In NLS:72, all baseline sample members

were spring term 1972 high school seniors. In HS&B, all members of the student sample were spring term 1980 sophomores or seniors. In ELS:2002, the base-year sample students were 10th-graders. Because NELS:88 base-year sample members were 8th-graders in 1988, its follow-ups encompass students (both in the modal grade progression sequence and out of sequence) and dropouts. Sample freshening was used in NELS:88 to provide cross-sectional nationally representative samples. Despite similarities, however, the sample designs of the four studies differ in three major ways: (1) the NELS:88 first and second follow-ups had relatively variable, small, and unrepresentative within-school student samples, compared to the relatively uniform, large, and representative within-school student samples in NLS:72 and HS&B; (2) unlike the two earlier studies, NELS:88 did not provide a nationally representative school sample in its follow-ups; and (3) there were differences in school and subgroup sampling and oversampling strategies in the four studies. These sample differences imply differences in the respondent populations covered by the four studies.

Questionnaire overlap is apparent among the four studies; nevertheless, caution is required when making trend comparisons. Some items were repeated in identical form across the studies; others appear to be essentially similar but have small differences in wording or response categories.

IRT scaling was used in the four studies to put math, vocabulary, and reading *test scores* on the same scale for 1972, 1980, 1982, and 2002 seniors. Additionally, there were common items in the HS&B and NELS:88 math tests that provide a basis for equating 1980–1990 and 1982–1992 math results, and common items in the NELS:88 and ELS:2002 reading and math tests that provide the link to obtain the ELS:2002 student ability estimates on the NELS:88 ability scale. In general, however, the tests in the four studies differed in many ways. Although group differences by standard deviation units may profitably be examined, caution should be exercised in drawing time-lag comparisons for cognitive test data.

Transcript studies in NELS:88, HS&B, ELS:2002, and NAEP were designed to support cross-cohort comparisons. The ELS:2002, NAEP, and NELS:88 studies, however, provide summary data in Carnegie units, whereas HS&B provides course totals. Note too that course offerings were only collected from schools that were part of the High School Effectiveness Study in the NELS:88 second follow-up, whereas course offerings were collected from *all* schools in HS&B (see chapter 7), and course offerings were collected from all base-year schools and the last school attended by

sample members who transferred out of their base-year school in ELS:2002 (see chapter 9).

Other factors should also be considered in assessing data compatibility. Differences in mode and time of survey administration across the cohorts may affect compatibility. NELS:88 seniors were generally surveyed earlier in the school year than were NLS:72 seniors. NLS:72 survey forms were administered by school personnel while HS&B and NELS:88 survey forms were administered primarily by contractor staff. There were also differences in questionnaire formats; the later tests had improved mapping and different answer sheets.

6. CONTACT INFORMATION

For content information on the NELS:88 project, contact:

Jeffrey Owings
Phone: (202) 502-7423
E-mail: jeffrey.owings@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Alt, M.N., and Bradby, D. (1999). *Procedures Guide for Transcript Studies* (NCES Working Paper 1999-05). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Haggerty, C., Dugoni, B., Reed, L., Cederlund, A., and Taylor, J. (1996). *National Education Longitudinal Study (NELS:88/94) Methodology Report* (NCES 96-174). National Center for Education Statistics. U.S. Department of Education. Washington, DC.

Horn, L., and West, J. (1992). *A Profile of Parents of Eighth Graders* (NCES 92-488). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S.J., Dowd, K.L., Stipe, J.L., Baldrige, J.D., Bartot, V.H., and Frankel, M.R. (1995). *User's Manual: NELS:88 Second Follow-Up: Dropout Component Data File* (NCES 93-375). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S.J., Scott, L.A., Taylor, J.R., Owings, J., and Quinn, P. (1998). *NELS:88 Base Year Through Second Follow-Up: Final Methodology Report* (NCES Working Paper 98-06). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Uses of Data

Curtin, T.R., Ingels, S.J., Wu, S., and Heuer, R. (2002). *User's Manual: NELS:88 Base-Year to Fourth Follow-up: Student Component Data File* (NCES 2002-323). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S., and Baldrige, J. (1995). *National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors* (NCES Working Paper 95-05). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S., and Dowd, K. (1995). *National Education Longitudinal Study of 1988: Conducting Trend Analyses: HS&B and NELS:88 Sophomore Cohort Dropouts* (NCES Working Paper 95-07). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Ingels, S., and Taylor, J. (1995). *National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data* (NCES Working Paper 95-06). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Survey Design

Bradby, D., and Hoachlander, G. (1999). *1998 Revision of the Secondary School Taxonomy* (NCES Working Paper 1999-06). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Quinn, P. (1995). *NELS:88 Second Follow-up Research and Development Working Papers* (NCES 94-251). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

- Ingels, S. (1996). *Sample Exclusion in NELS:88: Characteristics of Base Year Ineligible Students; Changes in Eligibility Status After Four Years* (NCES 96-723). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Rock, D.A., and Pollack, J.M. (1991). *Psychometric Report for the NELS:88 Base Year Test Battery* (NCES 91-468). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Spencer, B.D., Frankel, M.R., Ingels, S.J., Rasinski, K.A., and Tourangeau, R.E. (1990). *NELS:88 Base Year Sample Design Report* (NCES 90-463). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Wu, S., Levy, P.S., Wang, J., Chromy, J., and Li, L. (2003). *Report for Computation of Balanced Repeated Replicate (BRR) Weights for the Third (NELS88:1994) and Fourth (NELS88:2000) Follow-up Surveys* (NCES Working Paper 2003-18). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Data Quality and Comparability**
- Bokossa, M.C., and Huang, G.G. (2001). *Imputation of Test Scores in the National Education Longitudinal Study of 1988 (NELS:88)* (NCES Working Paper 2001-16). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Broene, P., and Rust, K. (2000). *Strengths and Limitations for Using SUDAAN, Stata, and WesVarPC for Computing Variances From NCES Data Sets* (NCES Working Paper 2000-03). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Burkam, D.T. (2003). *English Coursetaking and the NELS:88 Transcript Data* (NCES Working Paper 2003-02). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Burkam, D.T., and Lee, V. (2003). *Mathematics, Foreign Language, and Science Coursetaking and the NELS:88 Transcript Data* (NCES Working Paper 2003-01). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Kaufman, R., and Rasinski, K.A. (1991). *Quality of Responses of Eighth-Grade Students in NELS:88* (NCES 90-487). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- McLaughlin, D.H., Cohen, J., and Lee, R. (1997). *NELS:88 Survey Item Evaluation Report* (NCES 97-052). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Pollack, J.M., and Rock, D.A. (1997). *Constructed Response Tests in the NELS:88 High School Effectiveness Study* (NCES 97-804). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Rock, D.A., and Pollack, J.M. (1995). *Psychometric Report for the NELS:88 Base Year Through Second Follow-up* (NCES 95-382). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Salvucci, S., Walter, E., Colney, V., Fink, S., and Saba, M. (1997). *Measurement Error Studies at the National Center for Education Statistics* (NCES 97-464). National Center for Education Statistics, U.S. Department of Education. Washington, DC.
- Spencer, B.D., Frankel, M. R., Ingel, S.J., Rasinski, K.A., and Tourangeau, R. (1990). *NELS: 88 Base-Year Sample Design Report* (NCES 90-463). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Chapter 9: Education Longitudinal Study of 2002 (ELS:2002)

1. OVERVIEW

The Education Longitudinal Study of 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 of 2002 (the 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, switched to a homeschool environment, graduated early, or dropped out were administered a questionnaire. In the second follow-up in 2006, information was collected through a single electronic questionnaire about colleges applied to and aid offers received, enrollment in postsecondary education, employment and earnings, and living situation, including family formation. The third follow-up is planned for 2012, so that later outcomes, such as their persistence and attainment in higher education, or their transition into the labor market, can be understood in terms of their earlier aspirations, achievement, and high school experiences.

Purpose

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 and/or the world of work.

Components

ELS:2002 has two distinctive features. First, it is a longitudinal study in which the same units are surveyed repeatedly over time. Individual students will be followed for more than 10 years; the base-year schools were surveyed two times, once in 2002 and again in 2006. Second, in the high school years, it is an integrated multilevel study that involves multiple respondent populations. The respondents include students, their parents, their teachers, their librarians, and their schools.

Base-Year Survey. The base-year (2002) data collection instruments for 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).

LONGITUDINAL SAMPLE SURVEY OF THE 10TH-GRADE CLASS OF 2002; BASE-YEAR SURVEY, FIRST FOLLOW-UP IN 2004, AND SECOND FOLLOW-UP IN 2006

ELS:2002 collects data from:

- Students and dropouts
- School administrators
- Teachers
- Library media staff
- School facility checklist
- Parents
- High school transcripts

Student Questionnaire. 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. 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. Assessments in reading and mathematics were given at the same time. The baseline scores for the assessments can serve as a covariate or control variable for later analyses. Mathematics achievement was reassessed 2 years later, so that achievement gain over the last 2 years of high school could be measured and related to school processes and mathematics coursetaking.

Parent Questionnaire. One parent of each participating sophomore was asked to respond to a parent survey. The parent questionnaire was designed to gauge parents' aspirations for their child and to collect information about the home background and home education support system, the child's educational history prior to 10th grade, and parents' interactions with and opinions about the student's school.

Teacher Questionnaire. For each student enrolled in English or mathematics, a teacher was also selected to participate in a teacher survey. The teacher questionnaire was designed to illuminate questions on the quality, equality, and diversity of educational opportunity by obtaining information in two content areas: the teacher's evaluations of the student and information about the teacher's background and activities.

School Administrator Questionnaire. The school administrator questionnaire collected information on school characteristics, student characteristics, teaching staff characteristics, school policies and programs, technology, and 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 data from the school administrator questionnaire are nationally representative and can be used to generalize to the nation's regular high schools with sophomores in the 2001–02 school year.

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; the 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 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.

School Facilities Checklist. 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.

First Follow-up Survey. The first follow-up (2004) survey comprised seven questionnaires and an achievement test in mathematics. The questionnaires included a student questionnaire, a transfer student questionnaire, a new participant supplement questionnaire (NPSQ) (repeating selected questions from the base year), a homeschool student questionnaire, an early graduate questionnaire, a dropout (not currently in school) questionnaire, and a school administrator questionnaire.

Student questionnaire. 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 the 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. The full questionnaire comprised eight content modules: (1) contact information in support of the longitudinal design; (2) the student's school experiences and activities, including information about extracurricular participation, computer use in English and math, the transition process from the sophomore

year to upper-level secondary school, and the relationship of curricular programs and coursetaking to educational achievement and persistence; (3) time usage on homework, TV viewing, video and computer games, computers, nonschool reading, library utilization, and other activities; (4) plans and expectations for the future, including students' educational and life goals and values; (5) education after high school; (6) plans for work after high school; (7) work status and history; and (8) community, family, and friends.

Transfer student questionnaire. 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.

New participant supplement questionnaire (NPSQ). Any student new to the study at any of the core (base-year) schools was administered the NPSQ. 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. In addition, a subset of items included in the student questionnaire was also posed to new participants. These items (which are identical in content to those in 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 New Participant Supplement (NPS) gathered the key base-year variables that also were included in the NPSQ.

Homeschool student questionnaire. 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 the spring term of 2002 who were in a homeschool situation 2 years later. Homeschooled students were asked about their schooling activities and status, including their grade, coursework completed in science 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.

Early graduate questionnaire. Early graduates were defined as sophomore cohort members who had graduated from high school or received a General Educational Development (GED) credential on or before March 15, 2004. Early graduates completed only a subset of the items in 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.

Dropout questionnaire. 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 GED credential, and who had missed 4 or more consecutive weeks to a cause other than accident or illness. There was considerable overlap between the student and dropout questionnaires; both collected locating information for longitudinal follow-up and included items on school experiences and activities, time use, plans and expectations for the future, and the type and amount of work in which dropouts were engaged. The dropout questionnaire gathered information about students' work status and history, volunteer work or community college experience, and the educational behavior of friends. In the area of school experiences and activities, dropouts were asked questions about the school they last attended and their participation in alternative education 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.

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 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).

Second Follow-up Survey. The second follow-up (2006) survey was a single electronic questionnaire administered in three modalities—a web-enabled self-administration, computer-assisted telephone interviewing (CATI), and computer-assisted personal interviewing (CAPI). (Both CATI and CAPI are interviewer-administered modalities.) The questionnaire covered the transition from high school to postsecondary education, and included items on college access and choice. Items were drawn from a number of studies, including the Baccalaureate and Beyond Longitudinal Study (B&B, see chapter 16), Beginning Postsecondary Students (BPS, see chapter 15) Longitudinal Study, High School and Beyond (HS&B) Longitudinal Study (see chapter 7), National Education Longitudinal Study of 1988 (NELS:88, see chapter 8), and National Postsecondary Student Aid Study (NPSAS, see chapter 14). The interview was organized into four substantive sections: *High School*, *Postsecondary Education*, *Employment*, and *Community*. The interview concluded with a *Locating* section.

The first section, *High School*, collected retrospective information about high school completion. Respondents were classified as spring-term 2004 12th-graders, spring-term 2004 dropouts, neither, or both (for a small set). The majority of respondents skipped this section entirely because their high school completion date and the type of high school credential they earned were preloaded into the instrument at the start of data collection.

The *Postsecondary Education* section of the interview, the point of entry for most respondents, focused on education *after* high school. Questions pertained to the application process, admissions, financial aid offers, institutions attended, experiences at these institutions, and educational expectations. Complete month-by-month enrollment histories for all postsecondary institutions attended after high school were collected in this section. These enrollment histories (in conjunction with the date of high school completion or exit, as preloaded or reported in the *High School* section of the interview) were used to classify respondents into one of six mutually exclusive categories: standard enrollees, delayers, leavers, delayer-leavers, nonenrollees, and high school students. The questions administered to each respondent depended on his/her category. These categories were used for the *Employment* and *Community* sections as well. For more details, see the *Education Longitudinal Study of 2002: Base-Year to Second Follow-up Data File Documentation* (Ingels et al. 2007).

There were five topics in the *Employment* section. The questions for the first topic referred to the first job after high school. The second set of questions focused on employment at the time of the interview. The next set focused on jobs held by postsecondary students during the 2004–05 and 2005–06 academic years. Respondents were also questioned about months of unemployment (if a gap existed between high school and their first job, their first job and their current job, and/or their first job and the date of the interview, if they were not currently working). Lastly, the questions for the fifth topic focused on income, finances, and occupational expectations at age 30.

The final substantive section of the interview, *Community*, covered topics related to family formation, living arrangements, community involvement (including military service), and experiences that may influence the life course. With one minor exception, all questions pertained to all respondent types.

The interview concluded with the *Locating* section, which collected information that will be used to contact the respondents in the next round of the study.

High School Transcript Study. Transcripts were collected from sample members in late 2004 and early 2005, about 6 months to 1 year after most students had graduated from high school. Transcripts were collected from the students' base-year school. However, if it was learned during the first follow-up data collection that they had transferred, transcripts were collected from two schools: the base-year school and the last known school of attendance. For students who were added to the study during their senior year (known as "freshened" students), transcripts were only collected from their senior-year school. Transcripts were collected for regular graduates, as well as dropouts, early graduates, and students who were homeschooled after their sophomore year. For more information, see Chapter 29, High School Transcript (HST) Studies.

The ELS:2002 high school transcript data collection sought key pieces of information about coursetaking from students' official high school records (e.g., courses taken while attending secondary school, credits earned, year and term a specific course was taken, and final grades). When available, other information, such as dates enrolled, reason for leaving school, and standardized test scores, was collected. All information was transcribed and can be linked back to the students' questionnaire or assessment data. Because of the size and complexity of the file and the reporting variation by school, additional variables were constructed from the raw transcript file to facilitate analyses. These variables include standardized grade point averages

(GPAs), academic pipeline measures, and total credits earned by subject area. The construction of many of the transcript variables is based on Carnegie units. A Carnegie unit is equal to a course taken every day, one period per day, for a full school year.

Third Follow-up Survey. The third follow-up is planned for 2012. By this time, most of those who attended college will have graduated and entered the labor market. The third follow-up will collect data on the post-high school educational experiences of all sample members (such as their postsecondary persistence, attainment), their history of employment, family formation, community service, and other areas. Postsecondary transcripts will be obtained as well.

Periodicity

The base-year survey was conducted in the spring of 2002. The first follow-up was done in 2004, as was the high school transcript component. A post-high school follow-up was done in 2006. The third follow-up is planned for 2012; in this final follow-up, college transcripts will be obtained.

2. USES OF DATA

Using the multilevel and longitudinal information from 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 an 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 and/or the labor market. For students who continue on to higher education, data collected from the second follow-up and the third follow-up (which is planned for 2012) will help researchers measure the effects of these students' 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.

Apart from helping to describe the status of high school students and their schools, the second and third follow-up data 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, educational practitioners, and parents about the changes in the operation of the education 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—determinants and consequences;
- the role of family background and the home education support system in fostering students' educational success;
- the features of effective schools;
- the impact of coursetaking choices on success in the high school years (and thereafter);
- the equitable distribution of educational opportunities as registered in the distinctive school experiences and performance of students from various subgroups; and
- 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 using data from the second and third follow-ups. These issues include

- the later educational and labor market activities of high school dropouts;
- the transition of students who do not go directly on to postsecondary education or the world of work;
- access to, and choice of, undergraduate and graduate education 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.

3. KEY CONCEPTS

Cognitive Test Battery.

The test questions were selected from previous assessments: NELS:88, the National Assessment of Educational Progress (NAEP, see chapter 18), and Program for International Student Assessment (PISA, see chapter 22). Most, but not all, were multiple choice items. 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 were divided into process categories of skill/knowledge, understanding/comprehension, and problem solving. Through inclusion of items from the PISA, 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.

Cohort. A cohort is a group of individuals who have a statistical factor in common; for example, year of birth, grade in school, or year of high school graduation. ELS:2002 is a sophomore-grade cohort based on the spring term of the 2001–02 school year. It also contains, however, a nationally representative sample of high school seniors in the spring term of the 2003–04 school year.

Socioeconomic Status (SES). A composite variable is constructed through the combination of two or more variables—socioeconomic status, for example, combines mother’s education, father’s education, mother’s occupation, father’s occupation, and family income or an income proxy (household items) or it is calculated through the application of a mathematical function or transformation to a variable (e.g., conversion of raw test scores to percentile ranks).

Dropout. Dropouts were defined in ELS:2002 as sample members who had been absent from school for 4 or more consecutive weeks at the time of the survey and who were not absent due to accident or illness.

Early Graduate. Early graduates were defined as sample members who had graduated from high school or obtained certification of high school equivalency (e.g., obtained a GED credential) on or before March 15, 2004.

4. SURVEY DESIGN

Target Population

The ELS:2002 base year comprises two primary target populations—schools with 10th grades and 10th-grade students—in the spring term of the 2001–02 school year. There are two slightly different target populations for the first follow-up. One population consists of those students who were enrolled in the 10th grade in 2002. The other population consists of those students who were enrolled in the 12th grade in 2004. The former population includes students who dropped out of school between 10th and 12th grades, and such students are a major analytical subgroup. The target populations of the ELS:2002 second follow-up (2006) were the 2002 sophomore cohort and the 2004 senior cohort. The sophomore cohort consists of those students who were enrolled in the 10th grade in the spring of 2002 and the 12th-grade cohort comprises those students who were enrolled in the 12th grade in the spring of 2004. The sophomore cohort includes students who were in the 10th grade in 2002 but not in the 12th grade in 2004 (i.e., sophomore cohort members but not senior cohort members). The senior cohort includes students who were 12th-graders in 2004 but were not in the 10th grade in U.S. schools in 2002; they were included through a sample freshening process as part of the first follow-up activities.

Sample Design

The sample design for ELS:2002 is similar in many respects to the designs used in the three prior studies of the National Center for Education Statistics (NCES)

Longitudinal Studies Program: the National Longitudinal Study of the High School Class of 1972 (NLS:72), HS&B, and NELS:88. ELS:2002 is different from NELS:88 in that the ELS:2002 base-year sample students are 10th-graders rather than 8th-graders. As in NELS:88, there were oversamples of Hispanics and Asians in ELS:2002. However, for ELS:2002, counts of Hispanics and Asians were obtained from the Common Core of Data (CCD) and the Private School Universe Survey (PSS) to set the initial oversampling rates.

ELS:2002 used a two-stage sample selection process. First, schools were selected with probability proportional to size, and school contacting resulted in 1,220 eligible public, Catholic, and other private schools from a population of approximately 27,000 schools containing 10th-grade students. Of the eligible schools, 752 participated in the study. These schools were then asked to provide 10th-grade enrollment lists. In the second stage of sample selection, approximately 26 students per school were selected from these lists.

Base-Year Survey. The ELS:2002 base-year sample design comprises two primary target populations—schools with 10th grades and sophomores in these schools—in the spring term of the 2001–02 school year. The base-year survey used a two-stage sample selection process. First, schools were selected. These schools were then asked to provide sophomore enrollment lists.

The target population of schools for the ELS:2002 base year consisted of regular public schools, including state Department of Education schools and charter schools, and Catholic and other private schools that contained 10th grades and were in the United States (the 50 states and the District of Columbia). The sampling frame of schools was constructed with the intent to match the target population. However, selected schools were determined to be ineligible if they did not meet the definition of the target population. Responding schools were those schools that had a survey day (i.e., a day when data collection occurred for students in the school). Of the 1,270 sampled schools, there were 1,220 eligible schools and 752 responding schools (67.8 percent weighted response rate). School-level data reflect a school administrator questionnaire, a library media center questionnaire, a facilities checklist, and the aggregation of student data to the school level. School-level data, however, can also be reported at the student level and serve as contextual data for students.

The target population of students for the full-scale ELS:2002 consisted of spring-term sophomores in

2002 (excluding foreign exchange students) enrolled in schools in the school target population. The sampling frames of students within schools were constructed with the intent to match the target population. However, selected students were determined to be ineligible if they did not meet the definition of the target population. Of the 19,220 sampled students, there were 17,590 eligible students and 15,360 participants (87.3 percent weighted response rate). Student-level data consist of student questionnaire and assessment data and reports from students' teachers and parents.

First Follow-up Survey. The basis for the sampling frame for the first follow-up was the sample of schools and students used in the ELS:2002 base-year sample. There are two slightly different target populations for the follow-up. One population consists of those students who were enrolled in the 10th grade in 2002. The other population consists of those students who were enrolled in the 12th grade in 2004. The former population includes students who dropped out of school between 10th and 12th grades, and such students are a major analytical subgroup. Note that in the first follow-up, a student who is defined as a member of the student sample is either an ELS:2002 spring 2002 sophomore or a freshened first follow-up spring 2004 12th-grader.

If a base-year school split into two or more schools, many of the ELS base-year sample members moved en masse to a new school, and they were followed to the destination school. These schools can be thought of as additional base-year schools in a new form. Specifically, a necessary condition of adding a new school in the first follow-up was that it arose from a situation such as the splitting of an original base-year school, thus resulting in a large transfer of base-year sample members (usually to one school, but potentially to more). Four base-year schools split, and five new schools were spawned from these four schools. At these new schools, as well as at the original base-year schools, students were tested and interviewed. Additionally, student freshening was done, and the administrator questionnaire was administered.

Second Follow-up Survey. The target populations of the ELS:2002 second follow-up (2006) were the 2002 sophomore cohort and the 2004 senior cohort. The 2002 sophomore cohort consists of those students who were enrolled in the 10th grade in the spring of 2002, and the 2004 senior cohort comprises those students who were enrolled in the 12th grade in the spring of 2004. The sophomore cohort includes students enrolled in the 10th grade in 2002, but not in the 12th grade in 2004 (i.e., sophomore cohort members, but not senior

cohort members). The senior cohort includes students enrolled in the 12th grade in 2004, but not in the 10th grade in 2002; they were included through a sample freshening process as part of the first follow-up activities.

The second follow-up fielded sample consisted of 16,430 sample members: 14,100 respondents for both the base year and the first follow-up; 1,200 first follow-up nonrespondents who were base-year respondents; 650 base-year nonrespondents who were subsampled in the first follow-up and responded in the first follow-up; 210 base-year or first follow-up questionnaire-incapable members; 170 freshened respondents in the first follow-up; and 100 base-year respondents who were determined to be out of scope in the first follow-up. Once fielded, some members of the sample of 16,430 were determined to be out of scope. There were 460 out-of-scope second follow-up sample members who fell into five basic groups: deceased, out of country, institutionalized/incarcerated, questionnaire incapable/incapacitated, or unavailable for the duration of the 2006 data collection.

High School Transcript Study. Transcripts were collected for all sample members who participated in at least one of the first two student interviews: the base-year interview or the first follow-up interview. These sample members include base-year respondents who were first follow-up nonrespondents and base-year nonrespondents who were first follow-up respondents. Thus, sample members who were dropouts, freshened sample members, transfer students, homeschooled students, and early graduates are included if they were respondents in either of the first two student interviews. Transcripts were also requested for students who could not participate in either of the interviews because of a physical disability, a mental disability, or a language barrier.

Unlike previous NCES transcript studies, which collected transcripts only from the last school attended by sample members, the ELS:2002 transcript study collected transcripts from all base-year schools and the last school attended by sample members who transferred out of their base-year school. Incomplete records were obtained for sample members who had dropped out of school, had fallen behind the modal progression sequence, or were enrolled in a special education program requiring or allowing more than 12 years of schooling. Eighty-six percent of transcript respondents have 4 complete years of high school transcript information.

Data Collection and Processing

The base-year survey collected data from students, parents, teachers, librarians, and school administrators. Self-administered questionnaires and cognitive tests were the principal modes of data collection. Data collection took place primarily during in-school survey sessions conducted by Research Triangle Institute (RTI) field interviewer or team. Base-year data were collected in the spring term of the 2002 school year. A total of 752 high schools participated, resulting in a weighted school response rate of 67.8 percent. A total of 15,360 students participated, primarily in in-school sessions, for an 87.3 percent weighted response rate. Each sampled student's mathematics teacher and English teacher were given a questionnaire to complete. Weighted student-level coverage rates for teacher data were 91.6 percent (indicating receipt of a report from the math teacher, the English teacher, or both). School administrators and library media coordinators also completed a questionnaire (the weighted response rates were 98.5 percent and 95.9 percent, respectively). Questionnaires were mailed to parents, with a telephone follow-up for nonresponders. Student coverage for parent questionnaires was 87.5 percent (weighted). Survey administrators (SAs) completed a facilities checklist at each school. For the first follow-up, overall, about 89 percent (weighted) of the total ELS:2002 sample (comprising both 2002 sophomores 2 years later and 2004 freshened seniors) was successfully surveyed—whether through completion of a student, transfer student, dropout, homeschool, or early graduate questionnaire. For the second follow-up, the sample represents a subset of the combined population of 10th-graders in the spring term of 2002 and 12th-graders in the spring term of 2004. Of the total sample, approximately 15,900 were considered to be eligible for the 2006, among which 14,200 participated, resulting in a 88.4 weighted response rate.

Reference dates. In the base-year survey, most questions referred to the students' experience up to the time of the survey's administration in spring 2002. In the follow-ups, most questions referred to experiences that occurred between the previous survey and the current survey. For example, the first follow-up largely covered the period between 2002 (when the base-year survey was conducted) and 2004 (when the first follow-up was conducted).

Data collection. The base-year student data collection began in schools on January 21, 2002, and ended in schools in June 2002; telephone interviews with nonresponding students ended on August 4, 2002. Data collection from school administrators, library media center coordinators, and teachers ended in September

2002. The parent data collection ended on October 17, 2002. The first follow-up in-school data collection occurred between January and June 2004; out-of-school data collection took place between February and August 2004 and included telephone and in-person interviews. The second follow-up data collection was conducted from January to September 2006. To notify sample members about the start of data collection, all sample members and parent(s) were sent a packet which included instructions for the web-based survey.

During the field test of the base-year study, endorsements were secured from organizations felt to be influential in the eyes of the various entities being asked to participate (school administrators, librarians, teachers, students, and parents). Before school recruitment could begin, it was necessary to obtain permission to contact the schools. The Chief State School Officers (CSSOs) of each state (as well as the District of Columbia) were contacted to approve the study for the state. Permission to proceed to the district level was obtained in all 50 states as well as the District of Columbia. Once state approval was obtained, an information package was sent to the District Superintendent of each district/diocese that had sampled schools in the state. Permission to proceed to the school level was received from 693 of the 829 districts/dioceses having eligible sampled schools (83.6 percent). This represented a total of 891 eligible schools with district/diocese permission to be contacted among 1,060 eligible schools affiliated with districts/dioceses (84.1 percent). For public and Catholic schools, school-level contact was begun as soon as district/diocese approval was obtained. For private non-Catholic schools, it was not necessary to wait for higher approval, though endorsements from various private school organizations were sought. The principal of each cooperating school designated a school coordinator to serve as a point of contact at the school and to be responsible for handling the logistical arrangements. The coordinator was asked to provide an enrollment list of 10th-grade students. For each student, the coordinator was asked to give information about sex, race, and ethnicity, and whether the student had an Individualized Education Program (IEP). Dates for a survey day and two make-up days were scheduled. At the same time, staff members were designated to receive the school administrator and library media center questionnaires. Parental consents were obtained. On the survey day at each school, the survey administrator (SA) checked in with the school coordinator and collected any parental permission forms that had come in.

For the base-year and first follow-up surveys, the SA and survey administrator assistant (SAA) administered

the student questionnaire and tests via a group administration. The SA and SAA graded the routing tests (see details in the section of "Cognitive test data") and edited the student questionnaires for completeness. Makeup sessions were scheduled for students who were unable to attend the first session. Interviews were conducted by CATI for students who were unable to participate in the group-administered sessions. The school administrator, teacher, library media center, and parent questionnaires were self-administered; individuals who did not return their questionnaires by mail within a reasonable amount of time were followed up by telephone. The facilities checklist was completed by the SA based on his/her observations in the building on the school's survey day.

The first follow-up data collection required intensive tracing efforts to locate base-year sample members who, by 2004, were no longer in their 10th-grade schools, but had dispersed to many high schools. In the spring and again in the autumn of 2003, each base-year school was provided a list of ELS:2002 base-year sample members from their school. The school was asked to indicate whether each sample member was still enrolled at the school. For any sample member who was no longer enrolled, the school was asked to indicate the reason and date the student left. If the student had transferred to another school, the base-year school was asked to indicate the name and location of the transfer school. In the fall of 2003, each base-year school was also asked to provide a list of the 12th-graders enrolled at that school, so this information could be used in the freshening process. For students who had left their base-year school, the school was asked to provide contact information to allow for out-of-school data collection during the first follow-up survey period. Telephone data collection began in February 2004. Sample members identified for initial contact by the telephone unit included those no longer enrolled at the base-year school and those who attended base-year schools that did not grant permission to conduct an in-school survey session. Other cases were identified for telephone follow-up after the survey day and all makeup days had taken place at the school that the sample members attended. Some nonresponding sample members were assigned to SAs for field follow-up. A total of 797 sample members were interviewed in the field. An additional 80 field cases were completed either by mailed questionnaire or telephone interview and were withdrawn from the field assignment.

Data collection for the second follow-up was significantly redesigned to include survey modes and procedures that were completely independent of the in-school orientation of the first follow-up survey. An

important aspect of the second follow-up data collection was that high schools were no longer involved in providing assistance with locating sample members. Tracing and sampling maintenance techniques included the following: batch tracing services for updated address information and telephone numbers; updated locating information obtained from student federal financial aid applications; direct contact with sample members and their parents via mail, telephone, or the Internet; intensive tracing efforts by centralized tracing specialists; intensive tracing efforts by field locating specialists in local areas; and tracing students through postsecondary schools applied to or attended, as specified in the 2004 interview. Also, incentive payments were offered to respondents to maximize their participation.

There were three survey modes in the second follow-up: a web-enabled self-administered questionnaire, CATI, and CAPI. Data collection for the second follow-up began on January 25, 2006. For the first 4 weeks, only web and call-in data collection was made available to sample members. After the initial 4 weeks, outbound CATI data collection efforts were undertaken. The primary purpose of the CATI data collection was to complete telephone interviews with sample members when contacted or to set up an appointment to complete the interview. The CATI instrument was virtually identical to the web self-interview. (The only difference was that the CATI version provided an interviewer instruction on each screen to facilitate administration of each item.) CATI interviewers adhered to standardized interviewing techniques and other best practices in administering the interview. To reach sample members who had not yet participated by web or CATI modes, CAPI data collection commenced on April 17 (8 weeks after the start of outbound CATI calling). The approach for CAPI data collection followed the strategy used successfully in B&B:93/2003 and other recent NCES studies. This approach first identified geographic clusters according to the last known zip codes of sample members who could potentially be assigned to CAPI interviewing. Then, based on the distribution of cases by cluster, those that had the highest concentration of cases were staffed with one or more field interviewers. CAPI interviews were conducted on laptop computers via a web-based interface that used personal web server software. To maintain consistency across interviewing modes, the CAPI interview was identical to the CATI interview. CAPI interviewers were allowed to administer the interview over the telephone, which produced conditions even more similar to CATI interviewing.

Data processing. Data processing activities were quite similar for the base-year survey and the first follow-up. An initial check of student documents for missing data was performed on-site by the SA and SAA staff so that data could be retrieved from the students before they left the classroom. If a student neglected to answer a questionnaire item deemed to be critical, the SA/SAA asked the student to complete it after the end of the second-stage test (see details in the section of “Cognitive test data”).

All TELEform questionnaire scans were stored in an Structured Query Language (SQL) server database. CATI data were exported nightly to ASCII files. Cleaning programs were designed to concatenate CATI and TELEform SQL server data into SAS datasets, adjusting and cleaning variables when formats were not consistent. Special attention was focused on this concatenation to verify that results stayed consistent and to rule out possible format problems. Once questionnaire data were concatenated and cleaned across modes and versions, the following cleaning and editing steps were implemented:

- anomalous data cleaning based on a review of the data with the original questionnaire image;
- rule-based cleaning (changes that were made based on patterns in the data rather than on a review of the images);
- hard-coded edits based on changes recommended by a reviewer, if a respondent misunderstood the questionnaire (e.g., respondent was instructed to enter a percentage, but there was strong evidence that the respondent entered a count rather than a percentage); and
- edits based on logical patterns in the questionnaire (e.g., skip pattern relationships between gate and dependent questions).

All respondent records in the final dataset were verified with the Survey Control System (SCS) to spot inconsistencies. Furthermore, the data files served as a check against the SCS to ensure that all respondent information was included in production reports.

Data processing activities for the second follow-up differed from those in the base-year survey and the first follow-up, because respondents could complete a self-administered web questionnaire as an alternative to the survey modes used in previous years. A database was developed in which case/item-specific issues were

reviewed and new values were recorded for subsequent data cleaning and editing.

Item documentation procedures were developed in all waves of data collection to capture variable and value labels for each item. The wording of the question for each item was also provided as part of the documentation. This information was loaded into a documentation database that could export final data file layouts and format statements used to produce formatted frequencies for review. The documentation database also had tools to produce final electronic codebook input files.

Editing. An application was developed in which case/item-specific issues were reviewed and new values were recorded for subsequent data cleaning and editing. Records were selected for review based on one of the following criteria: random selection, suspicious values found during frequency reviews, values out of expected ranges, interviewer remarks, and values not adhering to a particular skip pattern. The review application provided the case/item-level information, the reason for the review, and a link to the scanned image of the questionnaire. Reviewers determined scanning corrections, recommended changes (if respondents had misinterpreted the question), and reviewed items randomly to spot potential problems that would require more widespread review.

The application was built on an SQL server database that contained all records for review and stored the recommended data changes. Editing programs built in SAS read the SQL server database to obtain the edits and applied the edits to the questionnaire data. Questionnaire data were stored at multiple stages across cleaning and editing programs, so comparison across each stage of data cleaning could be easily confirmed with the recommended edits. Raw data were never directly updated, so changes were always stored cumulatively and applied each time a cleaned dataset was produced. This process provided the ability to document all changes and easily fix errors or reverse decisions upon further review.

Editing programs also contained procedures that output inconsistent items across logical patterns within the questionnaire. For example, instructions to skip items could be based on previously answered questions; however, the respondent may not have followed the proper pattern based on the previous answers. These items were reviewed, and rules were written either to correct previously answered (or unanswered) questions to match the dependent items or blank out subsequent items to stay consistent with previously answered items.

Estimation Methods

The general purpose of the weighting scheme was to compensate for unequal probabilities of selection of students into the base-year sample and freshened students into the first follow-up sample and to adjust for the fact that not all students selected into the sample actually participated.

Weighting.

Student level. Two sets of student weights were computed. There is one set of weights for student questionnaire completion; this is the sole student weight that appears in the public-use file and generalizes to the population of spring 2002 sophomores who were capable of completing an ELS:2002 student questionnaire. A second set of weights, for the expanded sample of questionnaire-eligible and questionnaire-ineligible students, appears only in the restricted-use file. This weight sums to the total of all 10th-grade students.

First, the student-level design weight was calculated. The sample students were systematically selected from the enrollment lists at school-specific rates that were inversely proportional to the school's probability of selection. Specifically, the sampling rate for the student stratum within a school was calculated as the overall sampling rate divided by the school's probability of selection. To maintain control of the sample size and to accommodate in-school data collection, the sampling rates were adjusted, when necessary, so that no more than 35 students were selected. A minimum sample size constraint of 10 students was also imposed, if a school had more than 10 tenth-graders. Adjustments to the sampling rates were also made, as sampling progressed, to increase the sample size in certain student strata that were falling short of the sample size targets. The student sampling weight then was calculated as the reciprocal of the school-specific student sampling rate. The student nonresponse adjustment was performed using Generalized Exponential Models (GEMs) to compute the two student nonresponse adjustment factors. For data known for most, but not all, students, the data collected from responding students and weighted hot-deck imputation were used so that there would be data for all eligible sample students.

School level. School weights were computed in several steps. First, a school-level design weight equal to the reciprocal of the school's probability of selection was calculated; second, the school's design weight was adjusted to account for field-test sampling; third, the school weight was adjusted to account for the probability of the school being released. Next, GEMs, which are a unified approach to nonresponse

adjustment, poststratification, and extreme weight reduction, were used. For data known for most, but not all, schools that would be useful to include in the nonresponse adjustment, weighted hot-deck imputation was used so that there would be data for all eligible sample schools.

Scaling. Item Response Theory (IRT) was used to calibrate item parameters for all cognitive items administered to all students. This makes it possible to obtain scores on the same scale for students who took harder or easier forms of the test. IRT also permits vertical scaling of the two grade levels (10th grade in 2002 and 12th grade in 2004). A scale score estimating achievement level was assigned based on the pattern of right, wrong, and omitted responses on all items administered to an individual student. IRT postulates that the probability of correct responses to a set of test questions is a function of true proficiency and of one or more parameters specific to each test question. Rather than merely counting right and wrong responses, the IRT procedure also considers characteristics of each of the test items, such as their difficulty and the likelihood that they could be guessed correctly by low-ability individuals. IRT scores are less likely than simple number-right or formula scores to be distorted by correct guesses on difficult items if a student's response vector also contains incorrect answers to easier questions.

Imputation. In the base-year study, after the editing process (which included logical imputations), the remaining missing values for 14 analysis variables and two ability estimates (reading and mathematics) were statistically imputed. In the first follow-up study, two new variables were selected for imputation: the spring 2004 student ability estimate for mathematics and the spring 2004 student enrollment status. These variables were chosen because they are key variables used in standard reporting and cross-sectional estimation. Most of the variables were imputed using a weighted hot-deck procedure. Additionally, multiple imputations were used for a few variables, including test scores.

5. DATA QUALITY AND COMPARABILITY

Sampling Error

The variance estimation procedure had to take into account the complex sample design, including stratification and clustering. One common procedure for estimating variances of survey statistics is the Taylor series linearization procedure. This procedure takes the first-order Taylor series approximation of the nonlinear statistic and then substitutes the linear

representation into the appropriate variance formula based on the sample design. For stratified multistage surveys, the Taylor series procedure requires analysis strata and analysis primary sampling units (PSUs). Therefore, analysis strata and analysis PSUs were created. The impact of the departure of the ELS:2002 complex sample design from a simple random sample design on the precision of sample estimates can be measured by the design effect.

Design effects. The ELS:2002 sample departs from the assumption of simple random sampling in three major respects: student samples were stratified by student characteristics, students were selected with unequal probabilities of selection, and the sample of students was clustered by school. A simple random sample is, by contrast, unclustered and not stratified. Additionally, in a simple random sample, all members of the population have the same probability of selection. Generally, clustering and unequal probabilities of selection increase the variance of sample estimates relative to a simple random sample, and stratification decreases the variance of estimates.

In the ELS:2002 base-year study, standard errors and design effects were computed at the first stage (school level) and at the second stage (student level). The school administrator questionnaire was the basis for the school-level calculations; however, two items from the library questionnaire were also included. For student-level calculations, items from both the student and parent questionnaires were used. Therefore, three sets of standard errors and design effects were computed (school, student, and parent), which is similar to what was done for NELS:88. Each of the three sets includes standard errors and design effects for 30 means and proportions overall and for subgroups.

The student-level base-year design effects indicate that the ELS:2002 base-year sample was more efficient than the NELS:88 sample and the HS&B sample. For means and proportions based on student questionnaire data for all students, the average design effect in ELS:2002 was 2.35; the comparable figures were 3.86 for NELS:88 sophomores and 2.88 for the HS&B sophomore cohort. For all subgroups, the ELS:2002 design effects are smaller, on average, than those for the HS&B sophomore cohort. The smaller design effects in ELS:2002 compared to those for NELS:88 sophomores are probably due to disproportional strata representation introduced by subsampling in the NELS:88 first follow-up. The smaller design effects in ELS:2002 compared to those for the HS&B sophomore cohort may reflect the somewhat smaller cluster size used in the later survey. The ELS:2002 parent-level design effects are similar to the student-level design

effects. For estimates applying to all students, the average design effect was 2.24 for the parent data and 2.35 for the student data. For almost all subgroups, the average design effect was lower for the parent data than for the student data. The school-level design effects reflect only the impact of stratification and unequal probabilities of selection because the sample of schools was not clustered. Therefore, it could be expected that the design effects for estimates based on school data would be small compared to those for estimates based on student and parent data. However, this is not the case, as the school average design effect is 2.76. The reason for this is that the sample was designed to estimate students with low design effects. In addition to stratifying schools, a composite measure of size was used for school sample selection based on the number of students enrolled by race. This is different from the methodology used for NELS:88. The NELS:88 average school design effect in the base year study was considerably lower: 1.82.

The first follow-up design effects are lower for all respondents and for most of the subgroups than the base-year design effects. For the full sample, the design effect for males is the same as in the base year, the design effects for American Indian or Alaska Native and for multiracial respondents are greater than in the base year, and the design effects for the other 14 subgroups are lower than in the base year. For the panel sample, the design effects for American Indian or Alaska Native and for multiracial respondents are greater than in the base year, and the design effects for the other 15 subgroups are lower than in the base year.

The second follow-up design effects are lower for all respondents and for all of the common subgroups used in design effects calculations than the base-year and first follow-up design effects.

Nonsampling Error

Coverage error. In ELS:2002 base-year contextual samples, the coverage rate is the proportion of the responding student sample with a report from a given contextual source (e.g., the parent survey, the teacher survey, or the school administrator survey). For the teacher survey, the student coverage rate can be calculated as either the percentage of participating students with two teacher reports or the percentage with at least one teacher report. The teacher and parent surveys in ELS:2002 are purely contextual. The school-level surveys (school administrator, library media center, facilities checklist) can be used contextually (with the student as the unit of analysis) or in standalone fashion (with the school as the unit of analysis). Finally, test completions (reading assessments, mathematics assessments) are also

calculated on a base of the student questionnaire completers, rather than on the entire sample, and thus express a coverage rate. "Coverage" can also refer to the issue of missed target population units in the sampling frame (undercoverage) or duplicated or erroneously enumerated units (overcoverage).

Completed school administrator questionnaires provide 99.0 percent (weighted) coverage of all responding students. Completed library media center questionnaires provide 96.4 percent (weighted) coverage of all responding students. Of the 15,360 responding students, parent data (either by mailed questionnaire or by telephone interview) were received from 13,490 of their parents. This represents a weighted coverage rate of 87.4 percent.

Nonresponse error. Both unit nonresponse (nonparticipation in the survey by a sample member) and item nonresponse (missing value for a given questionnaire/test item) have been evaluated in ELS:2002.

Unit nonresponse. ELS:2002 has two levels of unit response (see table 6) : school response, defined as the school participating in the study by having a survey day on which the students took the test and completed the questionnaires; and student response, defined as a student completing at least a specified portion of the student questionnaire. The final overall school weighted response rate was 67.8 percent, and the final pool ¹ weighted response rate was 71.1 percent. The final student weighted response rate was 87.3 percent. Because the school response rate was less than 70 percent in some domains and overall, analyses were conducted to determine if school estimates were significantly biased due to nonresponse.

Nonresponding schools (or their districts) were asked to complete a school characteristics questionnaire. The nonresponding school questionnaire contained a subset of questions from the school administrator questionnaire that was completed by the principals of

¹ The sample was randomly divided by stratum into two release pools and a reserve pool. The two release pools were the basic sample, with the schools in the second pool being released randomly within stratum in waves as needed to achieve the sample size goal. Also, the reserve pool was released selectively in waves by simple random sampling within stratum for strata with low yield and/or response rates, when necessary. Each time schools were released from the second release pool or the reserve sample pool, sampling rates were adjusted to account for the non-responding schools and the new schools.

participating schools. (Of the 469 nonresponding eligible sample schools, a total of 437, or 93.2 percent, completed the special questionnaire.)

The school and student nonresponse bias analyses, in conjunction with the weighting adjustments, were not successful in eliminating all bias. However, they reduced bias and eliminated significant bias for the variables known for most respondents and nonrespondents, which were considered to be some of the more important classification and analysis variables. The relative bias decreased considerably after weight adjustments, especially when it was large before nonresponse adjustment, and the relative bias usually remained small after weight adjustments when it was small before nonresponse adjustment.

Student-level nonresponse. For students, although the overall weighted response rate was approximately 87 percent, the response rate was below 85 percent for certain domains, so a student-level nonresponse bias

typically relevant only for public schools). Consequently, only the school-supplied race/ethnicity and sex data, as well as the school-level data used in the school nonresponse bias analysis, were utilized in conducting the student-level nonresponse bias analysis.

For the student-level nonresponse bias analysis, the estimated bias decreased for every variable after weight adjustments were made. Therefore, the number of significantly biased variables decreased from 42 before adjustment to zero after adjustment.

Item nonresponse. There were no parent or teacher questionnaire items with a response rate that fell below 85 percent. However, there were 78 such items in the student questionnaire, including composites. Item nonresponse was an issue for the student questionnaire because, in timed sessions, not all students reached the final items. The highest nonresponse was seen in the final item, which was answered by only 64.6 percent of respondents.

Table 6. Unit-level and overall weighted response rates for selected ELS:2002 student populations, by data collection wave

Population	Unit-level weighted response rate			
	Base year school level	Base year student level	1st follow-up	2nd follow-up
Interviewed students	67.8	87.3	93.4	88.4
Tested students	67.8	95.1	87.4	†
Transfers	67.8	†	68.4	81.6
Dropouts	67.8	†	73.2	83.1
	Overall weighted response rate			
		Base year student level	1st follow-up	2nd follow-up
Interviewed students		59.2	63.3	59.9
Tested students		64.5	59.3	†
Transfers		†	46.4	55.3
Dropouts		†	49.6	56.3

† Not applicable.

SOURCE: Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2004). *ELS:2002 Base-Year Data File User's Manual* (NCES 2004-405). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2005). *Education Longitudinal Study of 2002/2004: Base-Year to First Follow-up Data File Documentation* (NCES 2006-344). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Ingels, S.J., Pratt, D.J., Wilson, D., Burns, L.J., Currihan, D., Rogers, J.E., and Hubbard-Bednasz, S. (2007). *Education Longitudinal Study of 2002: Base-Year to Second Follow-up Data File Documentation* (NCES 2008-347). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

analysis conditional on the school responding was also conducted. Some information on the characteristics of nonresponding students was available from student enrollment lists. On these lists, data were obtained on IEP status, race/ethnicity, and sex. These data were not provided by all schools (in particular, information on IEP status was often missing, and IEP information was

At the school level, 41 administrator items had a response rate that fell below 85 percent (ranging from a high of 84.7 percent to a low of 74.6 percent). No library media center questionnaire items fell below the 85 percent threshold, nor did any facility checklist items. While the school-level items can often be used as contextual data with the student as the basic unit of

analysis, these items are also, with the school weight, generalizable at the school level. Therefore, for the school administrator questionnaire, nonresponse rates and nonresponse bias estimates have been produced at the school level. While item nonresponse in the student questionnaire reflects item position in the questionnaire and the inability of some students to reach the final items in a timed session, nonresponse in the school questionnaire must be explained by two other factors: first, the nature of particular items; second, the fact that some administrators completed an abbreviated version of the questionnaire (the high nonresponse items did not appear in the abbreviated instrument).

Measurement error. In the field test, NCES evaluated measurement error in (1) student questionnaire data compared to parent questionnaire data; and (2) student cognitive test data. See *Education Longitudinal Study: 2002 Field Test Report* (Burns et al. 2003).

Parent-student convergence. Some questions were asked of both parents and students. This served two purposes: first, to assess the reliability of the information collected; second, to determine who was the better source for a given data element. These parallel items included number of siblings, use of a language other than English, and parent/child interactions. Additional items on parents' occupation and education, asked in both the parent and student interviews, were also evaluated for their reliability.

Parent-student convergence was low to medium, depending on the item. For example, the convergence on number of siblings is low. Although both parents and students were asked how many siblings the 10th-grader had, the questions were asked quite differently. It is not clear whether the high rate of disagreement is due to parents incorrectly including the 10th-grader in their count of siblings, the inaccurate reporting of "blended" families, or the differences in how the questions were asked in the two interviews. The parent-student convergence on parents' occupation and education was about 50 percent, very similar to those of the NELS:88 base-year interview.

Reliability of parent interview responses. In the field test, the temporal stability of a subset of items from the parent interview was evaluated through a reinterview administered to a randomly selected subsample of 147 respondents. The reinterview was designed to target items that were newly designed for the ELS:2002 interview or revised since their use in a prior NELS interview. Percent agreement and appropriate correlational analyses were used to estimate the response stability between the two interview administrations. The overall reliability of parent

interview responses varied from very high to very low, depending on the item. For example, the overall reliability for items pertaining to family composition and race and ethnicity is high; the overall reliability for items pertaining to religious background, parents' education, and educational expectations for the 10th-grader is only marginally acceptable.

Cognitive test data. The test questions were selected from previous assessments: NELS:88, NAEP, and PISA. Items were field tested 1 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. 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. 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.

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. In estimating a student's ability, IRT also accounts for each test question's difficulty, discriminating ability, and a guessing factor.

Data Comparability

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.

Comparability with NLS:72, HS&B, and NELS:88.

The ELS:2002 base-year and first follow-up surveys contained many data elements that were comparable to items from prior studies. *Some items are only approximate matches, and for these, analysts should judge whether they are sufficiently comparable for the analysis at hand. In other cases, question stems and response options correspond exactly across questionnaires.* These repeated items supply a basis for comparison with earlier sophomore cohorts (such as 1980 sophomores in HS&B and 1990 sophomores in NELS:88). With a freshened senior sample, the ELS:2002 first follow-up supports comparisons to 1972 (NLS:72), 1980 (HS&B), and 1992 (NELS:88). The first follow-up academic transcript component offers a further opportunity for cross-cohort comparisons with the high school transcript studies of HS&B, NELS:88, and NAEP.

Although the four studies have been designed to produce comparable results, they also have differences that may affect the comparability as well as the precision of estimates. Analysts should be aware of and take into account these several factors. In particular, there are differences in sample eligibility and sampling rates, in response rates, and in key classification variables, such as race and Hispanic ethnicity. Other differences (and possible threats to comparability) are imputation of missing data, differences in test content and reliability, differences in questionnaire content, potential mode effects in data collection, and possible questionnaire context and order effects.

Eligibility. Very similar definitions were used across the studies in deciding issues of school eligibility. Differences in student sampling eligibility, however, are more problematic. Although the target population is highly similar across the studies (all students who can validly be assessed or, at a minimum, meaningfully respond to the questionnaire), exclusion rules and their implementation have varied somewhat, and exclusion rates are known to differ, where they are known at all. For instance, a larger proportion of the student population was included in ELS:2002 (99 percent) than in NELS:88 (95 percent), which may affect cross-cohort estimates of change.

Sample design. Differences in sampling rates, sample sizes, and design effects across the studies also affect precision of estimation and comparability. Asian students, for example, were oversampled in NELS:88 and ELS:2002, but not in NLS:72 or HS&B, where their numbers were quite small. The base-year (1980) participating sample in HS&B numbered 30,030 sophomores. In contrast, 15,360 sophomores participated in the base year of ELS:2002. Cluster sizes

within school were much larger for HS&B (on average, 30 sophomores per school) than for ELS:2002 (just over 20 sophomores per school); larger cluster sizes are better for school effects research, but carry a penalty in greater sample inefficiency. Mean design effect (a measure of sample efficiency) is also quite variable across the studies: for example, for the 10th grade, it was 2.9 for HS&B and 3.9 for NELS:88 (reflecting high subsampling after the 8th-grade base year), with the most favorable design effect, 2.4, for the ELS:2002 base year. Other possible sources of difference between the cohorts that may impair change measurement are different levels of sample attrition over time and changes in the population of nonrespondents.

Imputation of missing data. One difference between the SES variable in ELS:2002 and in prior studies arises from the use of imputation in ELS:2002. Because all the constituents of SES are subject to imputation, it has been possible to create an SES composite with no missing data for ELS:2002. For the HS&B sophomores, SES was missing for around 9 percent of the participants, and for NELS:88 (in 1990) for just under 10 percent.

Score equating. ELS:2002 scores are reported on scales that permit comparisons with reading and mathematics data for NELS:88 10th-graders. 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 tests shared 30 reading and 49 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. Parameters for the common items were fixed at their NELS:88 values, resulting in parameter estimates for the noncommon items that were consistent with the NELS scale.

Transcript studies. ELS:2002, NELS:88, HS&B, and NAEP were designed to support cross-cohort comparisons. ELS:2002, NAEP, and NELS:88, however, provide summary data in Carnegie units, whereas HS&B provides course totals. In addition, unlike previous NCES transcript studies, which collected transcripts from the last school attended by the sample member, the ELS:2002 transcript study collected transcripts from all base-year schools and the last school attended by sample members who transferred out of their base-year school.

Other factors should be considered in assessing data compatibility. There are some mode-of-administration differences across the studies (for example, ELS:2002 collected 2006 data via self-administration on the Web, as well as by CATI and CAPI; in contrast, NLS:72 and

HS&B used paper-and pencil mail surveys). Order and context effects are also possible (questions have been added, dropped, and reordered, over time).

Comparability with PISA. A feature of ELS:2002 that expands its power beyond that of its predecessors is that it can be used to support international comparisons. Items from PISA were included in the ELS:2002 achievement tests. PISA, which is administered by the Organization for Economic Cooperation and Development, 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. ELS:2002 and PISA test instruments, scoring methods, and populations, however, differ in several respects that impact the equating procedures and interpretation of linked scores.

6. CONTACT INFORMATION

For content information on ELS:2002, contact:

John G. Wirt
Phone: (202) 502-7478
E-mail: john.wirt@ed.gov

Jeffrey A. Owings
Phone: (202) 502-7423
E-mail: jeffrey.owings@ed.gov

Mailing Address:

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

7. METHODOLOGY AND EVALUATION REPORTS

General

Bozick, R., Lytle, T., Sigel, P.H., Ingels, S.J., Rogers, J.E., Lauff, E., and Planty, M. (2006). *Education Longitudinal Study of 2002: First Follow-up Transcript Component Data File Documentation (Restricted Use)* (NCES 2006-338). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Cahalan, M.W., Ingels, S.J., Burns, L.J., Planty, M., and Daniel, B. (2006). *United States High School Sophomores: A Twenty-Two Year Comparison, 1980-2002* (NCES 2006-327). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Burns, L.J., Charleston, S., Chen, X., and Cataldi, E.F. (2005). *A Profile of the American High School Sophomore in 2002: Initial Results From the Base Year of the Longitudinal Study of 2002* (NCES 2005-338). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2004). *ELS: 2002 Base-Year Data File User's Manual* (NCES 2004-405). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2005). *Education Longitudinal Study of 2002/2004: Base-Year to First Follow-up Data File Documentation* (NCES 2006-344). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Wilson, D., Burns, L.J., Currivan, D., Rogers, J.E., and Hubbard-Bednasz, S. (2007). *Education Longitudinal Study of 2002: Base-Year to Second Follow-up Data File Documentation* (NCES 2008-347). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Uses of Data

Bozick, R., Lytle, T., Siegel, P.H., Ingels, S.J., Rogers, J.E., Lauff, E., and Planty, M. (2006). *Education Longitudinal Study: 2002/2004 Restricted-Use Base-Year, First Follow-up, and High School Transcript Data Files and Electronic Codebook System* (NCES 2006-351). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2004). *Education Longitudinal Study: 2002 Data Files and Electronic Codebook System* (NCES 2004-404). National Center for Education

Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2005). *Education Longitudinal Study: 2002/2004 Restricted-Use Base-Year to First Follow-up Data Files and Electronic Codebook System* (NCES 2006-330). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Rogers, J.E., Siegel, P.H., and Stutts, E. (2005). *Education Longitudinal Study:*

2002/2004 Public Use Base-Year to First Follow-up Data Files and Electronic Codebook System (NCES 2006-346). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Data Quality and Comparability

Burns, L.J., Heuer, R., Ingels, S.J., Pollack, J., Pratt, D.J., Rock, D., Rogers, J., Scott, L.A., Siegel, P., and Stutts, E. (2003). *Education Longitudinal Study: 2002 Field Test Report* (NCES Working Paper 2003-03). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.