Chapter 28: Crime and Safety Surveys

The National Center for Education Statistics (NCES) conducts two surveys on a regular basis to collect data on school crime and safety: the School Crime Supplement (SCS) to the National Crime Victimization Survey (NCVS), a survey of students ages 12 through 18; and the School Survey on Crime and Safety (SSOCS), a survey of public schools and principals.

TWO CRIME AND SAFETY SURVEYS:

- School Crime Supplement
- School Survey on Crime and Safety

1. SCHOOL CRIME SUPPLEMENT (SCS)

Overview

The SCS is conducted on a biennial basis as a supplement to the NCVS, which is administered by the Bureau of Justice Statistics (BJS), U.S. Department of Justice, and conducted by the U.S. Census Bureau. The NCVS is an ongoing household survey that gathers information on the criminal victimization of household members age 12 and older. NCES and BJS jointly created the SCS to study the relationship between victimization at school and the school environment.

The SCS is designed to assist policymakers—as well as academic researchers and practitioners at the federal, state, and local levels—in making informed decisions concerning crime in schools. The SCS gathers data from nationally representative samples of students who are between the ages of 12 and 18 and who are enrolled in grades 6–12 in U.S. public or private schools. Prior to 2007, eligible sample members were those who had attended school at any time during the 6 months preceding the interview. In 2007, the questionnaire was changed to include students who attend school at any time during the school year.

The SCS asks students a number of questions about their experiences with, and perceptions of, crime and violence occurring inside their school, on school grounds, on the school bus, and from 2001 onward, going to or from school. The SCS contains questions not included in the NCVS, such as those on preventive measures employed by schools; students' participation in after-school activities; students' perceptions of school rules and the enforcement of these rules; the presence of weapons, drugs, alcohol, and gangs in school; student bullying and cyber-bullying; hate-related incidents; and students' attitudes related to the fear of victimization at school. The SCS was conducted in 1989, 1995, 1999, 2001, 2003, 2005, 2007, and 2009. Future administrations are planned at 2-year intervals in odd-numbered years.

Sample Design

Each month, the U.S. Census Bureau selects respondents for the NCVS using a "rotating panel" design. Households are selected into the sample using a stratified, multistage cluster design. In the first stage, the primary sampling units (PSUs), consisting of counties or groups of counties, are selected, and smaller areas, called Enumeration Districts (EDs), are selected within each sampled PSU. Large PSUs are included in the sample automatically and are considered to be self-representing strata since all of them are selected. The remaining PSUs (called non-self-representing because only a subset of them are selected) are combined into strata by grouping PSUs with similar geographic and demographic characteristics, as determined by the decennial census. Within each ED, clusters of four households, called segments, are selected. Across all EDs, sampled households are then divided into discrete groups (rotations), and all age-eligible individuals in the households

become part of the panel. Such a design ensures a self-weighting probability sample of housing units and group-quarter dwellings within each of the selected areas. ("Self-weighting" means that prior to any weighting adjustments, each sample housing unit had the same overall probability of being selected.)

To account for units built within each of the sample areas after the decennial census, a sample of permits issued for the construction of residential housing is drawn. Jurisdictions that do not issue building permits are sampled using small land-area segments. These supplementary procedures, though yielding a relatively small portion of the sample, enable persons living in housing units built after the decennial census to be properly represented.

In order to conduct field interviews for the NCVS, the sample of households is divided into six groups, or rotations. Each group of households is interviewed seven times—once every 6 months over a period of 3 years. Each rotation group is further divided into six panels. A different panel of households, corresponding to one-sixth of each rotation group, is interviewed each month during the 6-month period. Because the NCVS is continuous, newly constructed housing units are selected as described above, and assigned to rotation groups and panels for subsequent incorporation into the sample. A new rotation group enters the sample every 6 months, replacing a group phased out after 3 years. This type of rotation scheme is used to reduce the respondent burden that might result if households were to remain in the sample permanently. It should be noted that the data from the NCVS/SCS interviews obtained in the incoming rotation are included in the SCS data files. The incoming rotation was included in the NCVS data file only in 2007.

Once in the panel, NCVS interviews are conducted with all household members age 12 or older. After completion of the NCVS interview, an SCS interview is given to eligible household members. In order to be eligible for the SCS, students must be 12 through 18 years old, have attended school in grades 6 through 12 at some point during the school year, and not have been homeschooled during the school year. Persons who have dropped out of school, have been expelled or suspended from school, or are temporarily absent from school for any other reason, such as illness or vacation, are eligible as long as they attended school at any time during the school year. For the 1989 and 1995 SCS, 19-year-old household members were considered eligible for the SCS interview. Prior to the 2007 SCS, household members who were enrolled in school sometime during the previous 6 months prior to the interview were eligible.

Data Collection and Processing

In all SCS survey years, the SCS was conducted for a 6-month period from January through June in all households selected for the NCVS. Eligible respondents were asked the supplemental questions in the SCS only after completing their entire NCVS interview.

The 2007 SCS was fully automated; all interviews were conducted through computer-assisted personal interviewing (CAPI), where field representatives used questionnaires loaded into laptop computers to conduct interviews, which could be completed either in person (for the first and subsequent interviews, as circumstances called for) or by telephone. Two modes of data collection were used through the 2005 collection: (1) paper-and-pencil interviewing, which was conducted in person for the first NCVS/SCS interview; and (2) computer-assisted telephone interviewing (CATI), unless circumstances called for an in-person interview. There were 5,620 students who participated in the SCS in 2007; 6,300 in 2005; 7,150 in 2003; 8,370 in 2001; 8,400 in 1999; 9,730 in 1995; and 10,450 in 1989. The 2009 data have been collected but not yet released.

Interviewers are instructed to conduct interviews in privacy unless respondents specifically agree to permit others to be present. Most interviews are conducted over the telephone, and most questions require "yes" or "no" answers, thereby affording respondents a further measure of privacy. While efforts are made to assure that interviews about student experiences at school are conducted with the students themselves, interviews with proxy respondents are accepted under certain circumstances. These include interviews scheduled with a child between the ages of 12 and 13 where parents refuse to allow an interview with the child; interviews where the subject child is unavailable during the period of data collection; and interviews where the child is physically or emotionally unable to answer for him- or herself.

Weighting

The purpose of the SCS is to be able to make inferences about criminal victimization in the 12- to 18-year-old student population in the United States. Before such inferences can be drawn, it is important to adjust, or "weight," the sample of students to ensure it is similar to the entire population in this age group. The SCS weights are a combination of household-level and person-level adjustment factors. In the NCVS, adjustments are made to account for both household- and person-level noninterviews. Additional factors are then applied to reduce the variance of the estimate by correcting for the differences between the sample

distributions of age, race, and sex and the known population distributions of these characteristics. The resulting weights are assigned to all interviewed households and persons in the file.

A special weighting adjustment is then made for the SCS respondents, and noninterview adjustment factors are computed to adjust for SCS interview nonresponse. This noninterview factor is applied to the NCVS person-level weight for each SCS respondent. Prior to 2007, two weights were available in the SCS data file. The first SCS weight was to be used if producing NCVS estimates using only the continuing rotations. The second SCS weight was derived using the final NCVS person weight that was calculated for all interviewed persons in continuing and incoming rotations. In 2007, all rotations were used for both the SCS and NCVS.

Table 20. Unweighted household, student, and overall unit response rates for the School Crime Survey: 1989–2007

Year	Household response rate	Student response rate	Overall response rate
1989	97	87	84
1995	95	78	74
1999	94	78	73
2001	93	77	72
2003	92	70	64
2005	91	62	56
2007	90	58	53

SOURCE: Chandler, K.A., Chapman, C.D., Rand, M.R., and Taylor, B.M. (1998). *Students' Reports of School Crime: 1989 and 1995* (NCES 98-241/NCJ-169607). National Center for Education Statistics, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC. Dinkes, R., Kemp, J., and Baum, K. (2009). *Indicators of School Crime and Safety: 2008* (NCES 2009-022/NCJ-226343). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Imputation

Item response rates are generally high. Most items are answered by over 95 percent of all eligible respondents. No explicit imputation procedure is used to correct for item nonresponse.

Sampling Error

Standard errors of percentage and population counts are calculated using the Taylor series approximation method using PSU and strata variables from the 1995, 1999, 2001, 2003, 2005, and 2007 data sets.

Another way in which the standard errors can be calculated, and were calculated in 1989, is by using the generalized variance function (GVF) constant parameters. The GVF represents the curve fitted to the individual standard errors that are calculated using the jackknife repeated replication technique.

Coverage Error

The decennial census is used for sampling housing units in the NCVS. To account for units built since the census was taken, supplemental procedures are implemented. (See "Sample Design" above.) Coverage error in the NCVS (and SCS), if any, would result from coverage error in the census and the supplemental procedures.

Unit Nonresponse

Because interviews with students can only be completed after households have responded to the NCVS, the unit completion rate for the SCS reflects both the household interview completion rate and the student interview completion rate (see table 20). Thus, the overall unweighted SCS response rate is calculated by multiplying the household completion rate by the student completion rate.

Due to the low student response rates in 2005 and 2007, unit nonresponse bias analyses were commissioned. In 2007, the analysis of unit nonresponse bias found evidence of bias by race, household income, and urbanicity variables. Hispanic respondents had lower response rates than respondents from other races/ethnicities. Respondents from households with an income of $25,000 or more had higher response rates than those from households with incomes of less than $7,500. Respondents who live in urban areas had lower response rates than those who live in rural areas. However, when responding students were compared to the eligible NCVS sample, there were no measurable differences between the responding students and the eligible students, suggesting the nonresponse bias has little impact on the overall estimates.

The analysis of unit nonresponse bias in 2005 also found evidence of bias for the race, household income, and urbanicity variables. White, non-Hispanic and other, non-Hispanic respondents had higher response rates than Black, non-Hispanic and Hispanic respondents. Respondents from households with incomes of $35,000–49,999 and $50,000 or more had higher response rates than those from households with incomes of less than $7,500, $7,500–14,999, $15,000–

24,999, and $25,000–34,999. Respondents who live in urban areas had lower response rates than those who live in rural or suburban areas.

Item Nonresponse

Item response rates for the SCS have been high. In all administrations, most items were answered by over 95 percent of all eligible respondents, with a few exceptions. One notable exception was the household income question, which was answered by about 80 percent of all households in 2007; about 82 percent of all households in 2005; and about 83, 84, 86, 90, and 90 percent of all households in 2003, 2001, 1999, 1995, and 1989, respectively. Due to their sensitive nature, income and income-related questions typically have relatively lower response rates than other items.

Measurement Error

Measurement error can result from respondents' different understandings of what constitutes a crime, memory lapses, and reluctance or refusal to report incidents of victimization. A change in the screener procedure between 1989 and 1995 was designed to result in the reporting of more incidents of victimization, more detail on the types of crime, and presumably more accurate data in 1995 than in 1989. (See "Data Comparability" below for further explanation.) Differences in the questions asked in the NCVS and SCS, as well as the sequencing of questions (SCS after NCVS), might have also lead to better recall in the SCS in 1995.

Data Comparability

The SCS questionnaire has been modified in several ways since its inception, as has the larger NCVS. Users making comparisons of data across years should be aware of the changes detailed below and their impact on data comparability. In 1989 and 1995, respondents to the SCS were asked two separate sets of questions regarding personal victimization. The first set of questions was part of the main NCVS, and the second set was part of the SCS. When examining data from either 1989 or 1995, the following have an impact on the comparability of data on victimization: (1) differences between years in the wording of victimization items in the NCVS as well as the SCS questionnaires; and (2) differences between SCS and NCVS items collecting similar data.

NCVS design changes. The NCVS was redesigned in 1992. Changes to the NCVS screening procedure put in place in 1992 make comparisons of 1989 data with those from later years difficult.

Due to the redesign, the victimization screening procedure used in 1995 and later years was meant to elicit a more complete tally of victimization incidents than the one used in 1989. For instance, it specifically asked whether respondents had been raped or otherwise sexually assaulted, whereas the 1989 screener did not. See *Effects of the Redesign on Victimization Estimates* (Kindermann, Lynch, and Cantor 1997) for more details on this issue.

In 2003, in accordance with changes to the Office of Management and Budget's standards for the classification of federal data on race and ethnicity, the NCVS item on race/ethnicity was modified. A question on Hispanic origin is now followed by a question on race. The new race question allows the respondent to choose more than one race and delineates Asian as a separate category from Native Hawaiian or Other Pacific Islander. An analysis conducted by the Demographic Surveys Division at the U.S. Census Bureau showed that the new race question had very little impact on the aggregate racial distribution of NCVS respondents, with one exception: there was a 2-percentage-point decrease in the percentage of respondents who reported themselves as White. Due to changes in race/ethnicity categories, comparisons of race/ethnicity across years should be made with caution.

In 2007, three changes were made to the NCVS for budgetary reasons. First, the sample was reduced by 14 percent beginning in July 2007. Second, to offset the impact of sample reduction, first-time interviews, which are not traditionally used in the production of the NCVS estimates, were included. Since respondents tend to report more victimization during first-time interviews than in subsequent interviews (in part, because new respondents tend to recall events having taken place at a time that was more recent than when they actually occurred), weighting adjustments were used to counteract a possible upward bias in the survey estimates. Using first-time interviews helped to ensure that the overall sample size would remain consistent with that in previous years. Lastly, in July 2007, the use of CATI as an interview technique was discontinued, and interviewing was conducted using only CAPI. For more details, see *Criminal Victimization, 2007* (U.S. Department of Justice 2008).

SCS design changes. The SCS questionnaire wording has been modified in several ways since its inception. Modifications have included changes in the series of questions pertaining to "fear" and "avoidance" between all survey years, beginning in 1995; changes in the definition of "at school" in 2001; changes in the introduction to, definition of, and placement of the item about "gangs" in 2001; and expansion of the single "bullying" question to include a series of questions in

2005 and including the topic of cyber-bullying in 2007. For more details, see *Student Victimization in U.S. Schools: Results From the 2005 School Crime Supplement to the National Crime Victimization Survey* (Bauer et al. 2008) and *Indicators of School Crime and Safety: 2008* (Dinkes, Kemp, and Baum 2009).

In addition, the reference time period for the 2007 SCS was revised from "the last 6 months" to "this school year." The change in reference period resulted in a change in eligibility criteria for participation in the 2007 SCS to include household members between ages 12 and 18 who had attended school at any time during the school year instead of during the 6 months preceding the interview, as in earlier surveys. This change was largely based on feedback obtained from students ages 12 to 18 during cognitive laboratory evaluations conducted by the U.S. Census Bureau. These respondents revealed they were not being strict in their interpretation of the 6-month reference period and were responding based on their experiences during the entire school year. Analyses of 2007 SCS data showed that estimates from 2007 are comparable to those from previous years. No change in reference period was made for criminal victimizations reported in the main NCVS.

Comparisons with related surveys. NCVS/SCS data have been analyzed and reported in conjunction with several other surveys on crime, safety, and risk behaviors. (See *Indicators of School Crime and Safety, 2008* [Dinkes, Kemp, and Baum 2009].) These include both NCES and non-NCES surveys. There are four NCES surveys: the School Safety and Discipline Questionnaire of the 1993 National Household Education Survey; the Teacher Questionnaire (specifically, the teacher victimization items) of the 1993–94, 1999–2000, 2003–04, and 2007–08 Schools and Staffing Survey; the Fast Response Survey System's Principal/School Disciplinarian Survey, conducted periodically; and the School Survey on Crime and Safety (SSOCS), conducted in 1999–2000, 2003–04, 2005–06, and 2007-08.

The non-NCES surveys and studies include the Youth Risk Behavior Surveillance System (YRBSS), a national and state-level epidemiological surveillance system developed by the Centers for Disease Control and Prevention (CDC) to monitor the prevalence of youth behaviors that most influence health; the School Associated Violent Death Study (SAVD), a study developed by the CDC (in conjunction with the U.S. Departments of Education and Justice) to describe the epidemiology of school-associated violent death in the United States and identify potential risk factors for these deaths; the Supplementary Homicide Reports (SHR), a part of the Uniform Crime Reporting (UCR) program conducted by the Federal Bureau of Investigation to provide incident-level information on criminal homicides; and the Web-based Injury Statistics Query and Reporting System Fatal (WISQARS Fatal), which provides data on injury-related mortality collected by the CDC.

Readers should exercise caution when doing cross-survey analyses using these data. While some of the data were collected from universe surveys, most were collected from sample surveys. Also, some questions may appear the same across surveys when, in fact, they were asked of different populations of students, in different years, at different locations, and about experiences that occurred within different periods of time. Because of these variations in collection procedures, timing, phrasing of questions, and so forth, the results from the different sources are not strictly comparable.

Contact Information

For content information on the SCS, contact:

NCES

Monica Hill
Phone: (202) 502-7379
E-mail: monica.hill@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

BJS

Michael Rand
Phone: (202) 616-3494
E-mail: randm@ojp.usdoj.gov

Methodology and Evaluation Reports

The reports listed below were either published by the U.S. Department of Education, National Center for Education Statistics (indicated by an NCES number), by the U.S. Department of Justice, Bureau of Justice Statistics, or were jointly published. See the technical notes in each report for a discussion of methodology.

General

U.S. Department of Justice, Bureau of Justice Statistics. (2008). *Criminal Victimization, 2007* (NCJ-224390). U.S. Department of Justice. Washington, DC: Bureau of Justice Statistics.

Uses of Data

Addington, L.A., Ruddy, S,A., Miller, A.K., and DeVoe, J.F. (2002). *Are America's Schools Safe? Students Speak Out: 1999 School Crime Supplement* (NCES 2002-331). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Bauer, L., Guerino, P., Nolle, K.L., and Tang, S. (2008). *Student Victimization in U.S. Schools: Results From the 2005 School Crime Supplement to the National Crime Victimization Survey* (NCES 2009-306). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Chandler, K.A., Chapman, C.D., Rand, M.R., and Taylor, B.M. (1998). *Students' Reports of School Crime: 1989 and 1995* (NCES 98-241/NCJ-169607). National Center for Education Statistics, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

DeVoe, J.F., and Bauer, L. (2010). *Student Victimization in U.S. Schools: Results From the 2007 School Crime Supplement to the National Crime Victimization Survey* (NCES 2010-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

DeVoe, J.F., and Kaffenberger, S. (2005). *Student Reports of Bullying: Results From the 2001 School Crime Supplement to the National Crime Victimization Survey* (NCES 2005–310). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

DeVoe, J.F., Peter, K., Noonan, M., Snyder, T.D., and Baum, K. (2005). *Indicators of School Crime and Safety: 2005* (NCES 2006-001/NCJ-210697). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Cataldi, E.F., Kena, G., and Baum, K. (2006). *Indicators of School Crime and Safety: 2006* (NCES 2007-003/NCJ-214262). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Cataldi, E.F., and Lin-Kelly, W. (2007). *Indicators of School Crime and Safety: 2007* (NCES 2008-021/NCJ-219553). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Kemp, J., and Baum, K. (2009). *Indicators of School Crime and Safety: 2008* (NCES 2009-022/NCJ-226343). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Kemp, J., and Baum, K. (2009). *Indicators of School Crime and Safety: 2009* (NCES 2010–012/NCJ 228478). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Ruddy, S., Bauer, L., and Neiman, S. (2010). *A Profile of Criminal Incidents at School: Results From the 2003–05 National Crime Victimization Survey Crime Incident Report* (NCES 2010-318). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Survey Design

Kindermann, C., Lynch, J., and Cantor, D. (1997). *Effects of the Redesign on Victimization Estimates* (NCJ-164381). U.S. Department of Justice. Washington, DC: Bureau of Justice Statistics.

2. SCHOOL SURVEY ON CRIME AND SAFETY

Overview

he School Survey on Crime and Safety (SSOCS) collects extensive crime and safety data from principals and school administrators of public schools. The survey builds on an earlier survey on school crime and safety conducted in 1997 using the Fast Response Survey System (FRSS). SSOCS focuses on incidents of specific crimes and offenses and a variety of specific discipline issues in public schools. It also covers characteristics of school policies, school violence prevention programs and policies, and school characteristics that have been associated with school crime. The survey is conducted with nationally

representative samples of regular public primary, middle, high, and combined schools in the 50 states and the District of Columbia. The sample does not include special education, alternative, or vocational schools; schools in the U.S. outlying areas and Puerto Rico, overseas Department of Defense schools, newly closed schools, home schools, Bureau of Indian Education schools, nonregular schools, ungraded schools, and schools with a high grade of kindergarten or lower.

Purpose. To collect detailed information on crime and safety from the schools' perspective; and to provide estimates of school crime, discipline, disorder, programs, and policies.

Components. SSOCS consists of a single questionnaire that is completed by principals or the person most knowledgeable about crime and safety issues at the school. Sections of the SSOCS questionnaire are composed of items about specific topics, including school practices and programs, parent and community involvement at school, school security, staff training, limitations on crime prevention, frequency of crime and violence at school, number of incidents, disciplinary problems and actions, and school characteristics.

Periodicity. SSOCS is administered to public primary, middle, high, and combined school principals in the spring of even-numbered school years. SSOCS is administered at the end of the school year to allow principals to report the most complete information possible. SSOCS was first administered in the spring of the 1999–2000 school year (SSOCS:2000). It has since been administered in the spring of the 2003–04, 2005–06, 2007–08, and 2009–10 school years (SSOCS:2004, SSOCS:2006, SSOCS:2008, and SSOCS:2010). A sixth collection is planned for the 2011–12 school year.

Uses of Data

SSOCS provides school-level data on crime and safety on the frequency of violence, the nature of the school environment, and the characteristics of school violence prevention programs. Such national data are valuable to policymakers and researchers who need to know what policies and programs are in place, what the level of crime is and how it is changing, and what disciplinary actions schools are taking. Some of the topics that may be examined are the following:

- Frequency and types of crimes at schools, including homicide, rape, sexual battery, attacks with or without weapons, robbery, theft, and vandalism;
- Frequency and types of disciplinary actions such as expulsions, transfers, and suspensions for selected offenses;
- Perceptions of other disciplinary problems such as bullying, verbal abuse, and disorder in the classroom;
- School policies and programs concerning crime and safety; and
- Pervasiveness of student and teacher involvement in efforts that are intended to prevent or reduce school violence.

The survey data also support analyses of how these topics are related to each other and how they are related to various school characteristics.

Sample Design

A stratified sample design is used to select schools for SSOCS. The sampling frame for SSOCS is constructed from the NCES Common Core of Data (CCD) Public Elementary/Secondary School Universe data file. Only "regular" schools (i.e., excluding special education, alternative, or vocational schools; schools in other U.S. jurisdictions; and schools that teach only prekindergarten, kindergarten, or adult education) are eligible for SSOCS. A stratified sample of 3,370 public schools was selected for SSOCS:2000; 3,740 public schools for SSOCS:2004; 3,570 public schools for SSOCS:2006; 3,480 for SSOCS:2008; and 3,476 for SSOCS:2010.

The same general sample design is used for each SSOCS. For sample allocation purposes, strata are defined by instructional level, type of locale, and enrollment size. Black, Hispanic, and other race/ethnicity status, and region were used as sorting variables in the sample selection process for SSOCS:2000, SSOCS:2004, SSOCS:2006, and SSOCS:2008 to induce additional implicit stratification. Beginning with SSOCS:2010, percent White enrolment and region were used as sorting variables. The three explicit and two implicit stratification variables have been shown to be related to school crime and thus create meaningful strata for this survey. The sample is designed to provide reasonably precise cross-sectional estimates for selected subgroups of interest.

Although the same design was used to allocate the sample across strata for all administrations of SSOCS, the calculation of the total initial sample differed between SSOCS:2000 and later SSOCS administrations. Without the experience of prior

administrations, stratum response rates had to be estimated for SSOCS:2000 when determining the number of sample cases within each stratum. In contrast, later administrations took advantage of the lessons learned from the prior data collection and used the prior stratum response rates to determine the proper size of the initial sample.

Data Collection and Processing

The data collection phase consists of (1) a mailout/ mailback stage; and (2) a telephone follow-up stage.

Reference dates. Data for SSOCS are collected at the end of even-numbered school years to allow principals to report the most complete information possible. For example, data collected in 2000 pertain to the 1999–2000 school year.

Data collection. SSOCS is conducted as a mail survey with telephone follow-up. Advance letters and, in some cases, e-mails, are sent to sampled schools informing them that they have been selected for SSOCS and describing the survey. SSOCS questionnaires are mailed to administrators with a cover letter describing the importance of the survey and a brochure providing additional information about it.

Starting approximately 1-2 weeks after the first questionnaire mailing, follow-up telephone prompts are used to verify that the questionnaire was received and to encourage survey response. As an alternative to replying by mail, data are also accepted by fax submission and by telephone.

After the data collection ends, returned questionnaires are examined for quality and completeness using both manual and computerized edits. Key items are identified. Depending on the total number of items that have missing or problematic data, and on whether these items have been designated as key items, data quality issues are resolved by recontacting the respondents or by imputation.

Editing. The survey questionnaires are reviewed to match survey responses with the appropriate values to be entered. After the data are key-entered, they are run through a series of editing programs: first, to determine whether a returned questionnaire can be considered complete; subsequently, to check data for consistency, valid data value ranges, and skip patterns.

Weighting

Data are weighted to compensate for differential probabilities of selection and to adjust for the effects of nonresponse.

Sample weights allow inferences to be made about the population from which the sample units are drawn. Because of the complex nature of the SSOCS sample design, these weights are necessary to obtain population-based estimates, to minimize bias arising from differences between responding and nonresponding schools, and to calibrate the data to known population characteristics in a way that reduces sampling error.

An initial (base) weight is first determined within each stratum by calculating the ratio of the number of schools available in the sampling frame to the number of schools selected. Because some schools refuse to participate, the responding schools do not necessarily constitute a random sample of the schools in the stratum. In order to reduce the potential of bias from nonresponse, weighting classes are determined by using a statistical algorithm similar to CHAID (i.e., chi-square automatic interaction detector) to partition the sample such that schools within a weighting class are homogeneous with respect to their probability of responding. The predictor variables for the analysis are school instructional level; locale; region; enrollment size; percent enrollment of Black, Hispanic, and other race/ethnicity students (or percent White enrollment for SSOCS:2010 and beyond); student-to-teacher ratio; percentage of students eligible for free or reduced-price lunch; and number of full-time-equivalent teachers. When the number of responding schools in a class is small, the weighting class is combined with another to avoid the possibility of large weights. After combining the necessary classes, the base weights are adjusted to produce nonresponse- adjusted weights, so that the weighted distribution of the responding schools resembles the initial distribution of the total sample.

The nonresponse-adjusted weights are then poststratified to calibrate the sample to known population totals in order to reduce bias in the estimates due to undercoverage. Two-dimension margins are set up for the poststratification: (1) instructional level and school enrollment size; and (2) instructional level and locale. An iterative process, known as the raking ratio adjustment, brings the weights into agreement with the known control totals. To be effective, the variables that define the poststrata must be correlated with the outcome of interest (school crime, for example). All three variables—instructional level, school enrollment size, and locale—have been shown to be correlated with school crime (Miller 2004).

Imputation

Completed SSOCS surveys contain some level of item nonresponse after the conclusion of the data collection

phase. Imputation procedures were used to impute missing values of key items in SSOCS:2000 and missing values of all items in each subsequent SSOCS. All imputed values are flagged as such.

SSOCS:2000. In SSOCS:2000, only the key data items with missing data in the file were imputed. Depending on the type of data to be imputed and the extent of missing values, a number of techniques—including hot-deck imputation, hot-deck imputation with collapsed imputation cell, logical imputation, and mean imputation—were employed.

SSOCS:2004 and beyond. In subsequent collections, imputation procedures were used to create values for all questionnaire items with missing data. This procedural change from SSOCS:2000 was implemented because the analysis of incomplete datasets may cause different users to arrive at different conclusions, depending on how the missing data are treated. The imputation methods used in SSOCS:2004 and later surveys were tailored to the nature of each survey item. Four methods were used: aggregate proportions, logical, best match, and clerical.

Future Plans

NCES plans to conduct SSOCS every 2 years in order to provide continued updates on crime and safety in U.S. public schools. SSOCS will next be administered in the 2011–12 school year.

Sampling Error

The estimators of sampling variances for SSOCS statistics take the SSOCS complex sample design into account. Both replication and Taylor Series methods are used to estimate sampling errors in SSOCS.

SSOCS utilizes the jackknife replication method, which involves partitioning the entire sample into a set of groups (replicates) based on the actual sample design of the survey. Survey estimates can then be produced for each of the replicates by utilizing replicate weights that mimic the actual weighting procedures used in the full sample. The variation in the estimates computed for the replicates can then be used to estimate the sampling errors of the estimates for the full sample. A total of 50 replicate weights were defined for each SSOCS.

Another approach to the valid estimation of sampling errors for complex sample designs is to use a Taylor series approximation. To produce standard errors using a Taylor series program, two variables are required (to identify the stratum and the Primary Sampling Unit [PSU]). The stratum-level variable is the indicator of the variance estimation stratum from which the unit was selected. The PSU is an arbitrary numeric identification number for the unit within the stratum.

Unit Nonresponse

A response rate is the ratio of the number of completed questionnaires to the number of cases sampled and eligible to complete the survey. All of the response rates are weighted to account for different probabilities of selection. Schools that are determined to be ineligible to participate in the survey (e.g. special education, alternative, or vocational schools; schools in other U.S. jurisdictions; and schools that teach only prekindergarten, kindergarten, or adult education) are not included in the calculation of response rates. For SSOCS:2000, the weighted response rate was 70 percent and the final number of respondents was about 2,270. For SSOCS:2004, the weighted response rate was 77 percent and the final number of respondents was about 2,770. For SSOCS:2006, the weighted response rate was 81 percent and the final number of respondents was about 2,720. For SSOCS:2008, the weighted response rate was 77 percent and the final number of respondents was about 2,560. As of the date of this publication, response rates were not yet available for SSOCS:2010. (See table 21 for weighted unit response rates by selected characteristics.)

Nonresponse bias analyses were conducted to determine if substantial bias is introduced due to school nonresponse. In SSOCS:2000, a CHAID analysis was conducted to group table cells to efficiently adjust for nonresponse, and regression analysis was used to confirm the choice of variables that resulted from the CHAID analysis. The study found virtually no significant differences in the estimates when comparing the initial nonresponse adjustments and the additional adjustments that were adopted based on the CHAID analysis. This suggests that much of the variation in response rates was captured in the original sampling strata. The adjustments to the weights were retained, despite their small impact, based on theoretical considerations that suggest they should be effective in attenuating nonresponse biases for a broad range of statistics.

In the 2004, 2006, and 2008 SSOCS, a number of analyses compared nonresponding and responding schools. The base-weighted distributions of the eight sampling frame variables—instructional level; type oflocale; region; school enrollment size; percent Black, Hispanic, and other race/ethnicity enrollment; student-to-teacher ratio; percentage of students eligible for free or reduced-price lunch; and number of full-time-equivalent teachers—were compared for responding and nonresponding schools. Then the differences and the full sample, using the base sampling weight,

Table 21. Weighted unit response rates, by selected school characteristics: 2000, 2004, 2006, and 2008

School characteristics	2000	2004	2006	2008
Total	70.0	77.2	81.3	77.2
Instructional level				
Primary	69.0	76.5	83.0	77.0
Middle	69.7	75.5	79.9	77.0
High school	71.0	77.8	78.8	76.2
Combined	79.6	84.9	75.7	80.8
Enrollment size				
Less than 300	76.3	86.0	83.2	83.3
300–499	70.9	77.8	84.7	76.7
500–999	67.5	72.8	79.9	76.2
1,000 or more	61.1	71.1	72.5	68.6
Urbanicity[1]				
City	63.6	69.0	75.4	69.4
Suburb	67.5	72.5	80.3	73.1
Town	75.4	84.9	86.7	84.6
Rural	77.0	86.1	85.5	83.9
Percent Black, Hispanic, and other race/ethnicity				
Less than 5 percent/ missing[2]	77.8	85.9	89.5	84.3
5 to 19 percent	71.3	77.7	82.8	80.8
20 to 49 percent	65.4	75.8	79.3	76.7
50 percent or more	64.6	71.4	76.7	71.4
Region				
Northeast	64.1	71.7	78.0	69.5
Midwest	74.0	80.8	83.2	80.8
South	77.1	79.8	82.5	79.7
West	64.3	75.7	80.9	74.6

[1]Starting with SSOCS:2008, a 12-category urban-centric locale variable from the NCES Common Core of Data (CCD) file was used; it was collapsed into 4 categories: city, suburb, town, and rural. Prior SSOCS collections used an 8-category CCD variable collapsed into 4 categories: city, urban fringe, town, and rural. Therefore, caution should be exercised when making direct comparisons between the 2008 and prior SSOCS collections.
[2]Beginning in 2008, there was no missing data for the race/ethnicity variable. This variable was imputed prior to sampling.
SOURCE: Chaney, B., Chowdhury, S., Chu, A., Lee, J., and Wobus, P. (2004). *School Survey on Crime and Safety (SSOCS) 2000 Public-Use Data Files, User's Manual, and Detailed Data Documentation* (NCES 2004-306). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Guerino, P., Hurwitz, M.D., Kaffenberger, S.M., Hoaglin, D.C., and Burnaska, K. (2007). *2003–04 School Survey on Crime and Safety Data File User's Manual* (NCES 2007-335). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Neiman, S., and DeVoe, J.F. (2009). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2007–08* (NCES 2009-326). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Nolle, K.L., Guerino, P., and Dinkes, R. (2007). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2005–06* (NCES 2007-361). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

between the respondent sample, using the final weight, were examined with respect to all eight sampling frame variables. Generally, the differences were not significant, leading to the conclusion that nonresponse bias is not an issue.

Item Nonresponse

Generally, item response rates were quite high. Because a more extensive follow-up was conducted when nonresponse was present for key items, item response rates were often higher for key items than for other questionnaire items.

For the 2008 SSOCS, weighted item response rates for individual items within the questionnaire ranged from 72 to 100 percent.

Of the 241 subitems in the 2008 SSOCS questionnaire, only 13 had response rates below 85 percent, and a nonresponse bias analysis was conducted on these 13 items. The detected bias was not deemed problematic enough to suppress any items from the data file.

Contact Information

For content information on SSOCS, contact:

Monica Hill
Phone: (202) 502-7379
E-mail: monica.hill@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

Methodology and Evaluation Reports

General

Chaney, B., Chowdhury, S., Chu, A., Lee, J., and Wobus, P. (2003). *2000 School Survey on Crime and Safety: Detailed Data Documentation* (NCES 2004-307). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Chaney, B., Chowdhury, S., Chu, A., Lee, J., and Wobus, P. (2003). *2000 School Survey on Crime and Safety: Public-Use Data File User's Manual* (NCES 2004-308). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Chaney, B., Chowdhury, S., Chu, A., Lee, J., and Wobus, P. (2004). *School Survey on Crime and Safety (SSOCS) 2000 Public-Use Data Files, User's Manual, and Detailed Data Documentation* (NCES 2004-306). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Guerino, P., Hurwitz, M.D., Kaffenberger, S.M., Hoaglin, D.C., and Burnaska, K. (2007). *2003–04 School Survey on Crime and Safety Data File User's Manual* (NCES 2007-335). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Izrael, D., and deFriesse, F.A. (2007). *2003–04 School Survey on Crime and Safety Public-Use Data File Codebook* (NCES 2007-333). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ruddy, S.A., Neiman, S., Bauer, L., Swaim, N.L., Thomas, T.L., and Parmer, R.J. (2009). *2005-06 School Survey on Crime and Safety (SSOCS) Survey Documentation for Data Users* (NCES 2010-320). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ruddy, S.A., Neiman, S., Hryczaniuk, C.A., Thomas, .L., and Parmer, R.J. (2010). *2007–08 School Survey on Crime and Safety (SSOCS) Survey Documentation for Public-Use Data File Users* (NCES 2010-307). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Tonsager, K., Neiman, S., Hryczaniuk, C., and Guan, X. (2010). *School Survey on Crime and Safety: 2007–08, Public-Use Data File Codebook* (NCES 2010-334). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Wallace, E.C., Neiman, S., Foster, K.M., and Guan, X. (2009). *2005–06 School Survey on Crime and Safety (SSOCS) Public-Use Data File Codebook* (NCES 2009-312). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Uses of Data

DeVoe, J.F., Peter, K., Noonan, M., Snyder, T.D., and Baum, K. (2005). *Indicators of School Crime and Safety: 2005* (NCES 2006-001/NCJ-210697). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice

Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Cataldi, E.F., Kena, G., and Baum, K. (2006). *Indicators of School Crime and Safety: 2006* (NCES 2007-003/NCJ 214262). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Cataldi, E.F., and Lin-Kelly, W. (2007). *Indicators of School Crime and Safety: 2007* (NCES 2008-021/NCJ-219553). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Dinkes, R., Kemp, J., and Baum, K. (2009). *Indicators of School Crime and Safety: 2008* (NCES 2009–022/NCJ 226343). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Guerino, P., Hurwitz, M.D., Noonan, M.E., and Kaffenberger, S.M. (2006). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2003–04* (NCES 2007-302REV). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Miller, A.K. (2004). *Violence in U.S. Public Schools: 2000 School Survey on Crime and Safety* (NCES 2004-314REV). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Neiman, S., and DeVoe, J.F. (2009). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2007–08* (NCES 2009-326). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Nolle, K.L., Guerino, P., and Dinkes, R. (2007). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2005–06* (NCES 2007-361). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Chapter 29: High School Transcript (HST) Studies

The value of school transcripts as objective, reliable measures of crucial aspects of students' educational experiences is widely recognized. NCES high school transcript studies collect information that is contained on the student high school record—i.e., courses taken while attending secondary school; information on credits earned; year and term a specific course was taken; and, final grades. When available, information on class rank and standardized scores is also collected. Once collected, information (e.g., course name, credits earned, course grades) is transcribed and standardized (e.g., credits and credit hours standardized to a common metric) and can be linked back to the student's questionnaire or assessment data.

Transcripts include information that is considered to be the official and fixed record regarding student course taking behaviors. This information is considered to be more accurate than student self-report information and represents a record of courses taken by the student. This information can be used to examine course-taking patterns of students and to predict future education outcomes.

Since 1982, NCES has conducted nine high school transcript studies: six high school transcript studies (HSTS) associated with the National Assessment of Educational Progress (NAEP), and three high school transcript studies as part of the Longitudinal Studies Program. Some of the key terms related to high school transcript studies are defined below.

Advanced Placement (AP). The AP Program is designed to prepare students to take the advanced placement examinations given by the Educational Testing Service (ETS). Students who pass these tests may be given credit and/or be exempted from requirements in colleges and universities based on their scores. Colleges and universities make their own rules regarding what tests to accept and the scores needed for credit or exemptions.

Carnegie unit. A factor used to standardize all credits indicated on transcripts across the study. The Carnegie unit is a strictly time-based reference for measuring secondary school attainment used by American universities and colleges. A single Carnegie unit is equal to 120 hours of classroom time over the course of a year at the secondary American high school level. Strictly speaking, this breaks down into a single 1-hour meeting on each of 5 days per week for a total of 24 weeks per year. However, knowing that classes usually meet for 50 minutes yields a value of 30 weeks per year. A semester (one-half of a full year) earns 1/2 Carnegie unit.

Catalog. A document compiled by a school or a district listing all available courses that are offered by the school and a description of those courses. Curriculum specialists review catalogs and use them to determine the appropriate Classification of Secondary School Courses (CSSC) code for each course.

Classification of Secondary School Courses (CSSC). A coding system employed for the purpose of standardizing high school transcript transcripts. The CSSC is a modification of the Classification of Instructional Programs (CIP) code used for

NAEP HIGH SCHOOL TRANSCRIPT STUDIES:

- National Assessment of Educational Progress High School Transcript Studies

LONGITUDINAL SURVEYS HIGH SCHOOL TRANSCRIPT STUDIES:

- High School and Beyond Longitudinal Study High School Transcript Study
- National Education Longitudinal Study of 1988 High School Transcript Study
- Education Longitudinal Study of 2002 High School Transcript Study

classifying college courses and contains approximate 2,300 course codes. Each CSSC code contains six digits. The first two digits identify the main program area, the second two digits represent a subcategory of courses within the main program area, and the final two digits define the specific course. For example, for the CSSC code 400522, the first two digits (40) define the Physical Sciences program area, the middle two digits (05) define the Chemistry subcategory, and the final two digits (22) define the course Advanced Chemistry.

Course offerings file. An high school transcript study data file that provides a comprehensive list of the courses offered in the schools included in the study. A CSSC code is associated with each course title.

International Baccalaureate (IB). A nonprofit educational foundation program consisting of a comprehensive 2-year international curriculum that allows students to fulfill the requirements of their national or state education systems.

Secondary School Taxonomy. The framework initially used by high school transcript for analyzing transcript data. The taxonomy divides high school coursework into three distinct curricula: academic, vocational, and personal/other.

Taxonomy. The classification of items into larger categories. In high school transcript studies, the items are specific secondary school courses (e.g., composition, first-year algebra, AP biology, American government) that are classified into course subject categories, as organized according to the Secondary School Taxonomy (SST), which is based on course content and level.

Tests and Honors file. A data file providing a list of honors and standardized test results, including SAT and ACT scores, that are found in the transcripts.

Transcript. A student's secondary school record containing courses taken, grades, graduation status, and attendance. In addition, it often includes scores from assessments, such as the PSAT, SAT, ACT, and a list of honors.

Transcript file. A data file providing a complete list of all courses appearing in the transcripts of students sampled in the study.

1. NAEP High School Transcript Studies

Since 1982, NCES has conducted six high school transcript studies (HSTS) associated with the National Assessment of Educational Progress (NAEP). NAEP has collected transcript data in 1987, 1990, 1994, 1998, 2000, 2005 and 2009 (see chapter 18). The results for 2009 HSTS will be reported in winter of 2011. Since information on 2009 HSTS is not yet available, 2009 HSTS will not be included in some sections of this chapter.

Components

Conducted in conjunction with NAEP, the 2009, 2005, 2000, 1998, 1994, 1990 and 1987 HSTS collected information on course offerings and coursetaking patterns in the nation's schools. Transcript data can be used to show coursetaking patterns across years that may be associated with proficiency in subjects assessed by NAEP.

Transcripts were collected for twelfth-grade students who graduated high school by the end of the collection period. Most students also participated in the NAEP assessments earlier that same year. Specifically, the students included in the 2005 and 2000 HSTS participated in the NAEP twelfth-grade mathematics and science assessments in 2005 and 2000 respectively; the students included in the 1998 HSTS participated in the civics, reading, and writing assessments in 1998; the students included in the 1994 HSTS participated in the geography, reading, and U.S. history assessments in 1994; the students included in the 1990 HSTS participated in the mathematics, science, and reading assessments in 1990; the students included in the 1987 HSTS participated in the 1986 long-term trend NAEP assessments in mathematics and science.

Periodicity

High school transcript studies have been conducted by NCES in conjunction with the NAEP since 1982. NAEP has collected transcript data in 1987, 1990, 1994, 1998, 2000, 2005, and 2009.

Survey Design

Target Population

The target population for high school transcript studies conducted as part of longitudinal surveys included all students in public and private schools who participated in previous data collections. For example, the target population for the 2004 high school transcript study included students who been in-school sophomores in the 2001–02 school year, participated in both the base-year and first follow-up interviews, completed the

mathematics assessment in the base-year and first follow-up interviews, and had complete transcript information for the 2002–03 and 2003–04 school years. The 2004 high school transcript study included 14,710 of the originally selected sample members of ELS:2002 sophomores in the spring of 2002 who were respondents in both the base-year and first follow-up interviews.

Sample Design

The NAEP High School Transcript Studies were conducted using nearly identical methodologies and techniques. They include the 2005, 2000, 1998, 1994, 1990, and 1987 transcript studies.

The 2005 High School Transcript Study. The sample design for the 2005 HSTS was designed to achieve a nationally representative sample of public and private high school graduates in the class of 2005. For public schools, the HSTS sample was the 12th-grade public school sample for the 2005 NAEP mathematics and science assessments; that is, the HSTS sample included every eligible sampled 2005 NAEP 12th-grade public school that was contacted for the HSTS, whether or not they actually participated in the NAEP assessments. For private schools, the HSTS sample was a subsample from the 2005 NAEP 12th-grade private school sample for the mathematics and science assessments. This subsampling process was carried out because private schools were oversampled in the 2005 NAEP. For the HSTS, the sample design called for the private schools' sample size to be proportionate to their share of eligible students. Over 26,000 transcripts from graduates were collected for the 2005 HSTS from a sample of about 640 public schools and 80 private schools.

For NAEP-participating schools, only those that participated in the main NAEP mathematics and science assessments were eligible for the HSTS. Within these schools, the HSTS used the same NAEP mathematics and science student samples. For schools that were selected for NAEP but did not participate, graduates were randomly selected. Approximately 94 percent of the HSTS sampled students were enrolled in schools that also participated in the NAEP assessments. Around 63 percent of the participating HSTS students also participated in NAEP.

The 2000 High School Transcript Study. The 2000 HSTS school sample comprised all 320 12th-grade public schools and a subsample of the 620 12th-grade private schools selected for the 2000 NAEP. The objective of private school subsampling was to reverse the oversampling of private schools in the 2000 NAEP so that the private school students in the 2000 HSTS would be represented in proportion to their prevalence in the general 12th-grade student population. While in NAEP 2000, private schools were oversampled to meet explicit target sample sizes for reporting group in order to provide reliable NAEP estimates for such students; in HSTS 2000, however, the oversampling of private schools was reversed so that the private school students in HSTS were represented in proportion to their prevalence in the general 12th-grade student population.

Because sampling was performed in most high schools prior to graduation, not all sampled students were, in fact, graduates. Only graduates, however, were eligible for the transcript study. From the exit status of the students, it was determined that of the 23,440 students in the sample, 21,090 actually graduated by October 2000 and 2,360 did not. From the 21,090 graduates, 20,930 transcripts were collected and processed. That is, 99 percent of the transcripts of eligible students were obtained.

The 1998 High School Transcript Study. The 1998 HSTS sample is nationally representative at both the school and student levels. The sample was composed of schools selected for the NAEP main sample that had 12th-grade classes and were within the 58 primary sampling units (PSUs) selected for the HSTS study. A subsample of 320 schools was selected from the eligible NAEP sample, consisting of 270 public schools and 50 nonpublic schools. In order to maintain as many links as possible with NAEP scores, replacement schools that were used in NAEP were also asked to participate in the transcript study, as opposed to sampling the NAEP refusal schools. Of the 320 schools in the original sample, 260 participated, of which 230 cooperated with both NAEP and HSTS and maintained links between students' transcript and NAEP data.

A total of 28,760 students were selected for inclusion in the HSTS study. Of these, 27,180 students were from schools that maintained their NAEP administration schedules and were identified by their NAEP booklet numbers. Another 500 students were from schools that participated in NAEP but had lost the link between student names and NAEP booklet numbers, and 1,080 were from schools that did not participate in NAEP. Of the 28,760 students in the original sample, 25,250 were deemed eligible for the transcript study, and 24,220 transcripts were collected and processed.

The 1994 High School Transcript Study. The 1994 HSTS sample of schools was nationally representative of all high schools in the United States. A subsample of 330 public schools and 50 private schools was

drawn from the lists of eligible NAEP public and private schools. One of these schools had no 12th-grade students and was not included in the HSTS study. Of the 380 remaining schools, 340 participated in the 1994 HSTS. The student sample was representative of graduating seniors from each school. Only those students were included whose transcripts indicated that they had graduated between January 1, 1994, and November 21, 1994. Approximately 90 percent of students in the 1994 HSTS also participated in the 1994 NAEP. The remaining students were sampled specifically for the transcript study, either because their schools did not agree to participate in the 1994 NAEP or because the schools participated in NAEP but did not retain the lists linking NAEP IDs to student names. The 1994 HSTS also included special education students who were excluded from the 1994 NAEP. High school transcripts were collected for 25,500 students from an eligible sample of 26,050 students.

The 1990 High School Transcript Study. The sample of schools was nationally representative of schools with a grade 12 or having 17-year-old students. (Some 380 schools were selected for the sample; some of these had no 12th-grade students.) The sample of students was representative of graduating seniors from each school. These students attended 330 schools that had previously been sampled for the 1990 NAEP. Approximately three-fourths of the sampled students had participated in the 1990 NAEP assessments. The remaining students attended schools that did not participate in NAEP or did not retain the lists linking student names to NAEP IDs. As with the later HSTS, only schools with a 12th grade were included, and only students who graduated from high school in 1990 were included. The 1990 HSTS also included special education students who had been excluded from the 1990 NAEP. In spring 1991, transcripts were requested for 23,270 students who graduated from high school in 1990; 21,610 transcripts were received.

The 1987 High School Transcript Study. The 1987 HSTS was conducted in conjunction with the long-term trend NAEP assessment. The schools in the 1987 HSTS were a nationally representative sample of 500 secondary schools that had been selected for the 1986 long-term trend NAEP assessments. The 1987 HSTS student sample represented an augmented sample of 1986 NAEP participants who were enrolled in the 11th grade and/or were 17 years old in the 1985–86 school year and who successfully completed their graduation requirements prior to fall 1987. The HSTS study included (1) students who were selected and retained for the 1986 NAEP assessment; (2) students who were sampled for the 1986 NAEP but were deliberately excluded due to severe mental, physical, or linguistic barriers; and (3) all students with disabilities attending schools selected for the 1986 assessment. Four of the participating schools had no eligible students without disabilities. Of the 500 schools selected for the HSTS study, 430 participated. There were 35,180 graduates in the sample, for whom 34,140 transcripts were received.

Data Collection and Processing

Data collection. The data collection procedures of the 2005 HSTS are discussed to illustrate the process. NAEP field workers requested sample materials for the 2005 HSTS when they first went to a school as part of the 2005 NAEP, and they collected these materials when they returned to the school for sampling. The sample materials included a list of courses offered for each of 4 consecutive years from school year 2001–02 through 2004–05; a completed School Information Form (SIF); and three sample transcripts, one representing a student taking "regular" courses, one with honors courses, and one with special education courses. For those students who were selected to participate in NAEP but who were classified as either having disabilities (SD) or English language learners (ELL), an SD/LEP questionnaire was completed for these students by the person most knowledgeable about the student. A School Questionnaire—which asked for information about school, teacher, and home factors that might relate to student achievement—was completed by a school official (usually the principal) as part of NAEP.

The SIF collected information about the school in general, sources of information within the school, course description materials, graduation requirements, grading practices, and the format of the school transcripts as part of the HSTS data collection process for non-NAEP participating schools.

In schools that did not participate in NAEP, the field worker first selected a sample of students, then requested transcripts for those students and followed the procedures for NAEP participants for reviewing and shipping transcripts. The SIF was also completed, and course catalogs for the past 4 school years were collected. The information in the catalogs was documented by completing the Course Catalog Checklist. At this point, the procedure was different from the one used for schools that participated in NAEP. Rather than obtaining and annotating three example transcripts, the field worker used the Transcript Format Checklist to annotate three actual transcripts from among those that were collected.

In the non-NAEP participating schools, the process of generating a sample of students began when the school

produced a listing of all students who graduated from the 12th grade during the spring or summer of 2005. This list was requested during the preliminary call placed to the school when it was determined that the school would participate in the HSTS. The following information was collected for each student in the HSTS: exit status; sex; date of birth (month/year); race/ethnicity; whether the student had a disability; whether the student was classified as limited English proficient; whether the student was receiving Title I services; and whether the student was a participant in the National School Lunch Program. These data were collected either with the list of 2005 graduates or after sampling, depending on which procedure was easier for the school.

Data processing. Each of the courses entered on the transcripts were coded using the Classification of Secondary School Courses (CSSC). For all NAEP transcript studies, courses appearing on student transcripts were coded to indicate whether they were transfer courses, held off campus, honors or above grade-level, remedial or below grade-level, or designed for students with limited English proficiency and/or taught in a language other than English.

Credit and grade information reported on transcripts also needed to be standardized. Standardization of credit information was based on the Carnegie unit, defined as the number of credits a student received for a course taken every day, one period per day, for a full school year.

The Computer-Assisted Coding and Editing (CACE) system was designed specifically for coding high school catalogs. CACE has two major components: (1) a component for selecting and entering the most appropriate CSSC code and "flags" for each course in a catalog; and (2) a component for matching each entry appearing on a transcript with the appropriate course title in the corresponding school's list of course offerings.

Each stage of the data coding and entering process included measures to ensure the quality and consistency of data. Measures to maintain the quality of data entry on transcripts included 100 percent verification of data entry; review of all transcripts where the number of credits reported for a given year (or the total number of credits) was not indicative of the school's normal course load or graduation requirements; and reconciliation of transcript IDs with the list of HSTS-valid IDs. Catalog coding reliability was maintained by conducting reliability checks. At least 10 percent of each school's course offerings were reentered by an experienced coder and the results compared with those of the original coder. If less than 90 percent of the entries agreed, the catalog was completely reviewed and any necessary changes were made. Agreement of 90 percent or better was found for approximately 85 percent of the school catalogs during the first review.

An additional quality check took place when the CACE files for a school were converted to delivery format. Reports listing frequencies of occurrences that might indicate errors were sent to the curriculum specialist for review. Each file was then assigned a status of 1 for complete, 2 for errors in transcript entry, 3 for errors in catalog coding and associations, or 4 for computer errors. A file with a status of 2, 3, or 4 was returned to Computer-Assisted Data Entry (CADE) and CACE for correction, a new report was generated, and the report was again reviewed. This process was repeated until the file had a status of 1, indicating that it was complete and correct.

Estimation Methods

Weighting. The weighting procedures are similar across the HSTS studies associated with NAEP. Only the 2005 NAEP HSTS procedures are described below. (For details on weighting in the other NAEP HSTS studies, see the relevant technical manuals.)

Two types of weights were created in the 2005 HSTS:

- HSTS base weights for all students who participated in the 2005 HSTS—that is, for whom a transcript was received and coded; and
- HSTS-NAEP linked weights for students who participated in both the 2005 HSTS and the 2005 NAEP. Linked weights were computed separately for mathematics and science assessment students. Each assessment sample represents the full population, so each of the two sets of assessment-linked weights aggregate separately to the population totals.

In each set of weights, the final weight attached to an individual student record reflected two major aspects of the sample design and the population surveyed. The first component, the base weight, reflected the probability of selection in the sample (the product of the probability of selecting the PSU, the probability of selecting the school within the PSU, and the probability of selecting the student within the school). The second component resulted from the adjustment of the base weight to account for nonresponse within the sample and to ensure that the resulting survey estimates of certain characteristics (race/ethnicity, size of

community, and region) conformed to those known reliably from external sources.

The final HSTS student weights were constructed in five steps:

(1) The student base weights (or design unbiased weight) were constructed as the reciprocal of the overall probability of selection.

(2) School nonresponse factors were computed, adjusting for schools that did not participate in the HSTS study. For the linked weights, adjustment factors were assigned for each session type (writing/civics, reading, and civics trend). The school nonresponse factors for the linked weights were also slightly different from the corresponding HSTS student weight school nonresponse factors to account for schools that refused to participate in NAEP.

(3) Student nonresponse factors were computed, adjusting the weights of responding students to account for nonresponding students. Definitions of responding and nonresponding students differed for the HSTS weights and the linked weights.

(4) Student trimming factors were generated to reduce the mean squared error of the resulting estimates. Another purpose of the trimming was to protect against a small number of large weights from dominating the resulting estimates of small domains of interest.

(5) The final step was poststratification, the process of adjusting weights proportionally so that they aggregate within certain subpopulations to independent estimates of these subpopulation totals. These independent estimates were obtained from Current Population Survey (CPS) estimates for various student subgroups. As the CPS estimates are associated with smaller sampling errors, this adjustment should improve the quality of the weights.

The linked student weights were constructed in a parallel manner, with some differences (e.g., the student base weight incorporated a factor for assignment to NAEP assessments). The school nonresponse factors were also slightly different for the linked weights to account for schools that refused to participate in the NAEP assessments. In addition, an extra nonresponse factor was computed for the linked weights to adjust for students whose transcripts were included in the HSTS study but who were absent from (or refused to participate in) a NAEP assessment. The trimming and poststratification steps for the linked weights were similar to those for the HSTS weights, with some differences. The missing transcript adjustments for the linked weights were very similar to those computed for HSTS weights.

Imputation. Imputation was done for missing data in the 1994, 1998, 2000, and 2005 High School Transcript Studies conducted in conjunction with NAEP.
In the 1994, 1998, 2000, and 2005 HSTS, it was not possible to obtain a transcript for a small percentage of high school graduates. In addition, some transcripts were considered unusable, since the number of standardized credits shown on the transcript was less than the number of credits required to graduate by the school. An adjustment is necessary in the weights of high school graduates with transcripts to account for missing and unusable transcripts. To do this adjustment correctly, it is necessary to have the complete set of high school graduates, with or without transcripts. Students who did not graduate were not included in this adjustment, but they were retained in the process for poststratification. There are a few students, however, for whom no transcripts were received and whose graduation status was unknown. Among these students, a certain percentage was imputed as graduating, based on the overall percentages of high school graduates. The remaining students were imputed as nongraduating. The imputation process was a standard (random within class) hot-deck imputation. For each student with unknown graduation status, a "donor" was randomly selected (without replacement) from the set of all students with known graduation status from the same region, school type, race/ethnicity, age class, school, and sex, in hierarchical order. The two race/ethnicity categories were (1) White, Asian, or Pacific Islander; and (2) Black, Hispanic, American Indian, or other. There were two age classes (born before 10/79; born during or after 10/79). Each student with known graduation status in a cell could be used up to three times as a donor for a student in the same cell with unknown graduation status. If insufficient donors were available within the cell, donors were randomly selected from students in another cell with similar characteristics to the cell in question. At the least, a donor had to be from the same region, type of school, race category, and age category.

Data Quality and Comparability

Sampling Error

Because of the HSTS multistage design, jackknife repeated replication was used for variance estimation in transcript studies associated with NAEP.

Table 22. Unweighted response rates for all eligible High School Transcript Study schools and students in each study: Various years, 1987-2005

Year	School response rate	Student coverage rate
2005	82	84[1]
2000	81	99
1998	88	98
1994	90	98
1990	87	93
1987[2]	87	97

[1] Weighted response rate.
[2] The 1987 HSTS was conducted in conjunction with the long-term trend NAEP assessment.
SOURCE: *The 1990 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1990, 1987, and 1982 High School Graduates* (No. ED360375). ERIC Document Reproduction Service. Washington, DC. Legum, S., Caldwell, N., Davis B., Haynes, J., Hill, T.J., Litavecz, S., Rizzo, L., Rust, K., Vo, N., and Gorman, S. (1997). *The 1994 High School Transcript Study Technical Report (*NCES 97-262). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Roey, S., Caldwell, N., Rust, K., Hicks, L., Lee, J., Perkins, R., Blumstein, E., and Brown, J. (2005). *The 2000 High School Transcript Study User's Guide and Technical Report* (NCES 2005-483). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Shettle, C., Cubell, M., Hoover, K., Kastberg, D., Legum, S., Lyons, M., Perkins, R., Rizzo, L., Roey, S., and Sickles, D. (2005). *The 2005 High School Transcript Study: The 2005 High School Transcript Study User's Guide and Technical Report* (NCES 2009-480). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Thorne, J. (1989). *High School Transcript Study, 1987* (No. ED315450). ERIC Document Reproduction Service. Washington, DC.

In the 2005 HSTS, a set of 62 replicate weights was attached to each record, one for each replicate. Variance estimation was performed by repeating the estimate procedure 63 times, once using the original full set of sample weights and once each for the set of 62 replicate weights. The variability among replicate estimates was used to derive an approximately unbiased estimate of the sampling variance. This procedure was used to obtain sampling errors for a large number of variables for the whole population and for specified subgroups.

Nonsampling Error

Coverage error. As the transcript studies associated with NAEP attempted to collect high school transcripts for all students selected for the assessment, whether or not they participated, transcripts for these students are included in the transcript study. Students who did not meet the graduation requirements established were excluded. Students with special education diplomas, certificates of attendance, and certificates of completion were also excluded, as were students with zero English credits and students with fewer than 16 Carnegie units. Because the NAEP studies collected data on the characteristics of excluded students, undercoverage bias can be quantified. Also, these studies were more inclusive in their transcript components than in their test or questionnaire administration. (See the section of "Sample Design".) It is believed that NAEP transcript studies had no transcript undercoverage due to exclusion of certain students.

Nonresponse error.

Unit nonresponse. There is unit nonresponse at both the school and student levels in HSTS. Response rates are presented in table 22.

An unweighted 82 percent of schools participated in the 2005 NAEP transcript study, higher than the 81 percent in the 2000 HSTS, but lower than the participation rate in the other NAEP transcript studies. Response rates varied with the characteristics of the sample school. For example, in 2005, despite a moderate overall response rate, only 57 percent of nonpublic schools responded.

At the student level, transcripts were obtained for 84 percent of eligible students in the 2005 HSTS (weighted), which is lower than the student-level response rate in the other transcript studies conducted in conjunction with NAEP. The response rate in the 2000 HSTS, 99 percent (unweighted), was the highest achieved in all six transcript studies.

Data Comparability

*Comparability of target populations***.** The target population of the 1987 NAEP HSTS has special features that affect its comparability to that of target populations in the other HSTS studies. The 1987 sample originated in a within-school representative sample of the schools' juniors/17-year-olds (students

born between October 1, 1968, and September 30, 1969). However, subsequent transfers into the school were given no chance of selection into the study; this fact qualifies how representative the within-school sample is and leaves it close to, but not precisely, a sample of the high school graduating class of 1987.

The 1990 HSTS sample originated within the 1990 NAEP sample of seniors/17-year-olds, but is further restricted to the seniors who in fact graduated in calendar year 1990. As such, it provides a nationally representative sample of 1990 high school graduates. Subsequent NAEP transcript collections have adhered to sample definitions that identify an unequivocally representative sample of graduating seniors.

*Sample inclusion and exclusion***.** A second issue concerns student sample inclusion and exclusion, especially with respect to students with disabilities and English language learners. The NAEP assessments collected information from school records about special education students. In the 1987 HSTS, the sample included students who were sampled for the assessment but deliberately excluded from it, as well as students with disabilities attending schools selected for the assessment. Thus, transcripts were collected for students excluded from the NAEP test as well as from the test-eligible sample. NAEP has carefully documented excluded students and identifies those who received testing accommodations.

Contact information

For content information about the High School Transcript Studies conducted in conjunction with NAEP, contact

Janis Brown
Phone: (202) 502-7482
E-mail: janis.brown@nces.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

Methodology and evaluation reports

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.

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.

Thorne, J. *(*1989). *High School Transcript Study, 1987* (No. ED315450). ERIC Document Reproduction Service. Washington, DC.

Caldwell, N. (1993). *The 1990 High School Transcript Study Data File User's Manual* (No. ED361354). ERIC Document Reproduction Service. Washington, DC.

Legum, S. (1993). *The 1990 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1990, 1987, and 1982 High School Graduates* (No. ED360375). ERIC Document Reproduction Service. Washington, DC.

Legum, S., Caldwell, N., Davis, B., Haynes, J., Hill, T.J., Litavecz, S., Rizzo, L., Rust, K., Vo, N., and Gorman, S. (1997). *The 1994 High School Transcript Study Technical Report (*NCES 97-262). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Perkins, R., Kleiner, B., Roey, S., and Brown, J. (2004). *The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990-2000* (NCES 2004-455). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Roey, S., Caldwell, N., Rust, K., Blumstein, E., Krenzke, T., Legum, S., Kuhn, J., Waksberg, M., Haynes, J., and Brown, J. (2001). *The 1998 High School Transcript Study User's Guide and Technical Report* (NCES 2001-477). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Roey, S., Caldwell, N., Rust, K., Blumstein, E., Krenzke, T., Legum, S., Kuhn, J., Waksberg, M., Haynes, J., and Brown, J. (2001). *The 1998 High School Transcript Tabulations: Comparative Data on Credits Earned and Demographics for 1998, 1994, 1990, 1987, and 1982 High School Graduates* (NCES 2001-498). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Roey, S., Caldwell, N., Rust, K., Hicks, L., Lee, J., Perkins, R., Blumstein, E., and Brown, J. (2005). *The 2000 High School Transcript Study User's Guide and Technical Report* (NCES 2005-483). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Shettle, C., Cubell, M., Hoover, K., Kastberg, D., Legum, S., Lyons, M., Perkins, R., Rizzo, L., Roey, S., and Sickles, D. (2008) *The 2005 High School Transcript Study: The 2005 High School Transcript Study User's Guide and Technical Report.* (NCES 2009-480). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Shettle, C., Roey, S., Mordica, J., Perkins, R., Nord, C., Teodorovic, J., Brown, J., Lyons, M., Averett, C., and Kastberg, D. (2007). *America's High School Graduates: Results From the 2005 NAEP High School Transcript Study* (NCES 2007-467). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

2. Longitudinal Surveys High School Transcript Studies

Overview

Since 1982, NCES has conducted three high school transcript studies as part of the Longitudinal Studies Program. The first NCES-sponsored transcript study was conducted in 1982, as part of the first follow-up to the High School and Beyond (HS&B) Longitudinal Study (see chapter 7). In 1992, another transcript study was conducted in conjunction with the second follow-up to the National Education Longitudinal Study of 1988 (NELS:88) (see chapter 8). A third transcript study associated with the longitudinal study series was conducted in 2004, as part of the first follow-up to the Education Longitudinal Study of 2002 (ELS:2002) (see chapter 9).

Components

The 2004 High School Transcript Study. The ELS:2002 high school transcript data collection sought information about coursetaking from the student's official high school record—including courses taken while attending secondary school, information on credits earned, year and term a specific course was taken, and final grades. When available, other information was collected, including dates enrolled, reason for leaving school, and standardized test scores. Once collected, the information was transcribed and can be linked back to the student's questionnaire or assessment data collected by ELS. Due to the size and complexity of the file, and because of reporting variations by school, additional variables were constructed from the raw transcript file. These composite variables include standardized grade point average (GPA), high school academic program, total credits earned by subject, and others.

The 1992 High School Transcript Study. The NELS:1992 high school transcript data include detailed information about the types of degree programs, periods of enrollment, majors or fields of study, specific courses taken, grades and credits attained, and credentials earned.

The 1982 High School Transcript Study. The HS&B transcript data collection allows the study of the coursetaking behavior of the members of the 1980 sophomore cohort throughout their 4 years of high school. Data include 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 course was taken; GPA; days absent; and standardized test scores.

Periodicity

High school transcript studies have been conducted by NCES as part of the Longitudinal Studies Program since 1982. Transcript studies associated with the Longitudinal Studies Program were conducted in 1982, 1992, and 2004.

Survey Design

Target Population

The target population for high school transcript studies conducted as part of longitudinal surveys included all students in public and private schools who participated in previous data collections. For example, the target population for the 2004 high school transcript study included students who been in-school sophomores in the 2001–02 school year, participated in both the base-year and first follow-up interviews, completed the mathematics assessment in the base-year and first follow-up interviews, and had complete transcript information for the 2002–03 and 2003–04 school years. The 2004 high school transcript study included 14,710 of the originally selected sample members of ELS:2002 sophomores in the spring of 2002 who were respondents in both the base-year and first follow-up interviews.

Sample Design

Sample design is essentially similar across the various administrations of the high school transcript studies: multistage, stratified, and clustered design.

The 2004 ELS High School Transcript Study. This study was conducted as part of the ELS:2002 first follow-up in 2004 (see chapter 9). A total of 1,550 out of 1,950 schools participated in the request for

transcripts for an unweighted participation rate of 79 percent. The base-year school weighted response rate is 95 percent. The course offerings response rate for base-year schools is 88 percent. Ninety-one percent(91 percent, weighted) of the entire student sample have some transcript information (14,920 out of 16,370 students).

Transcripts were collected from the school that the students were originally sampled from in the base year (which was the only school for most sample members) and from their last school of attendance if it was learned during the first follow-up student data collection that they had transferred. 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. For freshened students, transcripts were only collected from their senior year school. Transcripts were collected for regular graduates, dropouts, early graduates, and students who were homeschooled after their sophomore year.

The 1992 High School Transcript Study. This transcript study was conducted as part of the NELS:88 second follow-up (see chapter 8). A total of 2,260 schools were identified as longitudinal cohort eligible for the high school transcript study, in the second follow-up tracing of the NELS:88 first follow-up sample. Since the high school transcript study conducted as part of NELS:88 was limited to 1,500 schools for the full range of data (student, parent, teacher, school administrator, and transcript data) collection, it was necessary to select a sample of schools. All schools identified as having four or more first follow-up sample members enrolled were included in the school-level sample with certainty (probability = 1.0), and random samples were selected for retention from schools identified as having three first follow-up members (probability = 0.75), two first follow-up members (probability = 0.65), and one first follow-up member (probability = 0.31845). (Note that by the time of the data collection, only 1,380 of the 1,500 schools contained at least one NELS sample member.) Transcript and contextual data were requested for all students in the 1,380 selected schools.

In addition, transcripts were collected for all dropouts, early graduates, and 12th-grade sample members ineligible for the base-year, first follow-up, and second follow-up surveys owing to a language, physical, or mental barrier (triple ineligibles), through the sample "freshening" process, and the followback process of excluded students. This added 470 schools to the sample.

Of the 1,840 schools in the 1992 sample (including both contextual[1] and noncontextual schools), 1,540 participated in the 1992 study. Transcripts were requested for 19,320 students, and 17,290 transcripts were received.

The 1982 High School Transcript Study. The first transcript study was a component of the HS&B first follow-up. The 1982 study included students from 1,900 secondary schools—1,000 HS&B sampled schools and 900 schools to which students selected for the transcript survey had transferred (and for which no data collection activities other than transcript collection were carried out). Of these 1,900 schools, 1,720 provided transcripts. The total student sample size was 18,430 students. From the 1980 sophomores selected for the HS&B first follow-up, 12,310 cases were retained in the study sample with certainty—12,030 cases in the probability sample plus 280 nonsampled co-twins. In addition, a systematic sample of 6,120 cases was subsampled from the 17,700 remaining first follow-up selections, with a uniform probability of approximately .35. Transcripts were collected for 15,940 of the 18,430 students.

Data Collection and Processing

Data collection. The data collection and processing procedures are similar across the three transcript studies conducted as part of the Longitudinal Studies Program. The data collection procedures of the 2004 high school transcript study are discussed to illustrate the data collection process.

The ELS:2002 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. Collecting the transcripts in the 2004–05 school year allowed for more complete high school records. Transcripts were collected from the school that the students were originally sampled from in the base year (which was the only school for most sample members) and from their last school of attendance, if it was learned during the first follow-up student data collection that they had transferred. By requesting transcripts and related information for transfer students from a second school, this ELS:2002 transcript study offers the unique advantage of having extensive information on multiple school attendance and, therefore, increased accuracy of enrollmen histories. 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

[1] Schools selected for the contextual components of the second follow-up-the school administrator and teacher surveys-are referred to as contextual schools.

special education program requiring or allowing more than 12 years of schooling. For freshened students, transcripts were only collected from their senior year school. Transcripts were collected for regular graduates, dropouts, early graduates, and students who were homeschooled after their sophomore year.

From December 2004 through June 2005, survey materials were sent to over 2,000 schools. This group included schools that participated either in the base-year or first follow-up survey and transfer schools that were first contacted regarding ELS:2002 during transcript data collection. Transcripts were not requested from 10 base-year schools because they had refused to participate in the first follow-up survey. Additionally, transcripts were not requested from one base-year school that had no eligible students. Schools were paid $5 for each transcript. Transcripts were requested for over 16,000 sample members. Included were sample members who were ineligible to participate in the base year or first follow-up because of a physical disability, a mental disability, or a language barrier. Ninety-five schools required explicit consent from sample members or their parents/guardians before releasing transcript information. Of the sample members who attended these schools, about a quarter provided signed release forms. Two weeks after the survey materials were sent to the school, a follow-up postcard was sent as a reminder to complete the data collection forms and to send the requested materials to the Research Triangle Institute (RTI). If, after an additional week, RTI had not received the materials from the school, assigned institutional contactors (ICs) began telephone prompting to request that the materials be sent as soon as possible. Nonresponding schools contacted during the telephone prompting frequently requested remailing of the data collection materials. During telephone contacts, the ICs also identified any additional requirements the school had for releasing transcripts. Telephone follow-up with schools continued through June 2005. Additional measures were implemented to ensure an adequate response rate. In June 2005, data collection materials were sent to schools that had not yet provided all of the requested transcripts. In addition, in-person visits to nonresponding schools were conducted during April through June 2005 to collect the requested materials or to assist the school transcript preparer in assembling the information. For efficiency, the schools were selected for in-person visits by their proximity to other schools. In-person visits were made only to schools that had not sent transcript materials for any requested sample members.

Data processing. Each of the courses entered on the transcripts were coded using the Classification of Secondary School Courses (CSSC). The descriptions of the 2004 high school transcript data processing procedures illustrate the data processing done in the three transcript studies conducted as part of the Longitudinal Studies Program.

For the 2004 data processing, incoming data collection forms, transcripts, and course catalogs were logged into the survey control system by staff from RTI. Course catalog and transcript data were then entered using a web-based CADE system. Course catalogs from ELS:2002 base-year schools were keyed and coded for the preparation of course offerings data. For ELS:2002 base-year schools that provided them, courses listed in course catalogs were keyed and assigned the appropriate CSSC code before transcript keying and coding. For each catalog course entered, keyer-coders selected an appropriate course code from the CSSC look-up table in the data entry system. All transcripts received from a school were assigned to a single person for keying and coding. Course catalogs from non-base-year schools were not keyed. Data entry of each catalog and transcript was reviewed for accuracy by a supervisor or by a group of keyer-coders trained to perform these reviews. Procedures for editing, coding, error resolution, and documentation were modeled after the NELS:88 second follow-up transcript component (Ingels et al. 1995). Data entry systems included checks for valid variable ranges and codes, including legitimate missing codes, and CSSC code checks. Sequences of machine edits and visual data inspections were performed. Tasks included supplying missing data, detecting and correcting illegal codes, and investigating and resolving inconsistencies or anomalies in the data. Variable frequencies and cross-tabulations were reviewed to verify the correctness of machine editing.

Estimation Methods

Weighting. The weighting procedures used in the 2004 high school transcript study are presented as an example of the weighting procedures used in transcript studies conducted as part of the Longitudinal Studies Program.

In the 2004 high school transcript study, weight was assigned as follows. First, the first follow-up design weight was used as the starting weight. Next, Generalized Exponential Models (GEM) were used to compute weight adjustments. Weight adjustments included (1) a nonresponse adjustment to reduce potential bias owing to transcript nonresponse; and (2) a poststratification adjustment to ensure that sums for weights for certain domains had the same totals as those in the first follow-up. The nonresponse adjustment was performed in two stages: (1) at the

school refusal stage (e.g., the school refused to provide any transcript); and (2) at the within-school student-level nonresponse stage (see below for more details). Poststratification was performed to keep key estimates consistent with those in the first follow-up. Extreme weights were adjusted, truncated, and smoothed by GEM as part of the nonresponse and poststratification adjustments rather than as a separate step.

Imputation. Imputation was done for missing data in the High School Transcript Studies conducted for NELS and HS&B as part of the Longitudinal Studies Program.

Imputation was done for missing sex data in the 1992 NELS transcript study, using the student's first name to determine sex. In the 1982 HS&B transcript study, values were imputed for missing sex and race/ethnicity.

Data Quality and Comparability

Sampling Error

For the 1982, 1992, and 2004 high school transcript studies, variance estimation required the Taylor series linearization procedure, which took into account the complex sample design of these surveys, including stratification and clustering. 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 PSUs (in ELS:2002, schools are the PSUs). Therefore, analysis strata and analysis PSUs were created in the base year and used again in the first follow-up.

Transcript studies conducted as part of the Longitudinal Studies Program may also use the Balanced Repeated Replication (BRR) variance estimation procedure or both Taylor Series linearization and BRR for variance estimation. For example, in NELS:88 and ELS:2002, variance estimation can be done in two ways: first, with Taylor Series linearization using software such as SUDAAN, AM, or STATA when using the Electronic Codebook (ECB) data; or, when using BRR, using the table generator (DAS—Data Analysis System) version of the dataset. Thus, the same estimate can have two different standard errors even within the same study, depending on whether its basis is a Taylor Series linearization or BRR. HS&B used both BRR and the Taylor Series and compared the results. These two methods result in very small differences that should not markedly change conclusions about the standard error of an estimate.

Coverage error. Potential sources of undercoverage in the high school transcript studies include (1) incomplete sampling frame data, as no national listing of schools is, or remains for very long, 100 percent complete and accurate; (2) omissions and errors in school rosters; and (3) deliberate exclusion of certain categories of students—such as students with physical or mental disabilities or non-English speakers, who might find it difficult or impossible to complete demanding cognitive tests and questionnaires. The first two sources are thought to have only a very small impact on high school transcript estimates. The most serious potential source is the undercoverage bias due to the exclusion of certain categories of students.

HS&B and NELS transcript studies are believed to exclude students with physical, mental, or linguistic barriers to assessment or survey participation. NELS transcript study collected data on the characteristics of excluded students, so that undercoverage bias can be quantified, and that the 1992 NELS study had negligible undercoverage of about 3 percent for the senior cohort. Although quantifiable exclusion data are not available for HS&B, given the similarity of eligibility rules in all two studies, it is reasonable to presume that HS&B exclusion rates were between 3 and 6 percent.

Nonresponse error.

Unit nonresponse. There is unit nonresponse at both the school and student levels in high school transcript studies. Response rates for all eligible high school transcript schools and students are presented in table 23.

Table 23. Unweighted response rates for all eligible High School Transcript Study schools and students in each study: 1982, 1992 an 2004

Year	School response rate	Student coverage rate
2004	79	91
1992	84	89
1982	91	88

SOURCE: Bozick, R., Lyttle, T., Siegel, 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 (*NCES 2006-338). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Transcripts were collected from 79 percent (unweighted) of the schools in the ELS:2004 study, 84 percent (unweighted) of the schools in the NELS:92 study, and 91 percent (unweighted) of the schools in the 1982 HS&B study.

At the student level, transcripts were obtained for 91 percent (unweighted) of eligible students in the ELS: 2004 study, 89 percent in the NELS:92 study, and 88 percent (unweighted) in the 1982 HS&B study.

Item nonresponse. Rates for item nonresponse have ranged from nonexistent to extremely high, depending on the type of item, across all of the high school transcript studies. As would be expected in transcript studies, course-level items have little if any nonresponse. Specific items include school year, term, and grade in which a course was taken; school-assigned course credits; and standardized course grade. However, nonresponse rates for items such as class size, cumulative GPA, class rank, days absent in each of the 4 high school years, and standardized test scores (e.g., PSAT, SAT, ACT) are very high.

In the 1992 NELS transcript study, the nonresponse rates for these items ranged from 0 percent for school year to less than 2 percent for the school term in which a course was taken. Incompleteness of actual course data, while considered to be limited, is another source of potential bias in a transcript study. Course data may be incomplete for students who transferred from one school to another. Also, it is difficult to assess the completeness of transcript data for dropouts in the 1982 HS&B and 1992 NELS transcript studies because of inconsistencies between enrollment reports of the sample member and the school.

Transcripts often provide other pieces of information that are useful in the analysis of coursetaking patterns: days absent in each school year, class rank, class size, month and year student left school, reason student left school (e.g., dropped out, graduated, transferred), cumulative GPA, participation in specialized courses or programs, and various standardized test scores (e.g., PSAT, SAT, ACT). While nonresponse rates for participation in specialized courses or programs (2 percent) and month/ year/reason student left school (less than 4 percent) are quite low in the 1992 NELS transcript study, nonresponse rates for the other items are very high: 18 percent for class size; 22 percent for cumulative GPA; 23 percent for class rank; 42–44 percent for days absent in each of the 4 high school years; and 67–73 percent for standardized test scores. (Note that although students were asked in a student questionnaire whether and when they planned to take specific tests, some students may not have actually taken the tests; this would, in part, explain the high nonresponse rates for test scores.

This wide range of item nonresponse rates is comparable to the range of nonresponse rates in the 1982 HS&B transcript study. For example, in the 1982 HS&B transcript study, the nonresponse rate was 32 percent for class rank and class size, 41–47 percent for days absent per school year, and 75 percent and above for standardized test scores.

Two key analytic variables are sex and race/ethnicity. Item nonresponse rates for sex have been extremely low: 0 percent in both the 1982 HS&B transcript study and the 1992 NELS transcript study. For race/ethnicity, nonresponse has ranged from 0 percent in the 1982 HS&B transcript study to 0.7 percent in the 1992 NELS transcript study.

Measurement error. Possible sources of measurement error in high school transcript studies are differences between schools and teachers in grading practices (e.g., grade inflation), differences in how data are recorded (although efforts are made to standardize grades and course credits for the high school transcript studies), and errors in keying or processing the transcript data (although the system has many built-in quality checks). The amount of measurement error in any survey or study is difficult to determine, and it is unknown for the high school transcript studies. However, because the transcripts are official school records of students' progress, it is reasonable to presume that there is less measurement error than in other types of data collections, particularly those that are self-reported.

Data Comparability

The high school transcript studies conducted by NCES have both similarities and dissimilarities of design and methodology that raise questions of comparability and may sometimes require analytical adjustments to ensure that comparability is maximized. This section presents four such issues: the comparability of target populations, sample inclusion and exclusion, methodology across studies, and content across studies. For details, please refer to the *Education Longitudinal Study of 2002: First Follow-up Transcript Component Data File Documentation* (Bozick et al. 2006).

*Comparability of target populations***.** The first comparability issue concerns the comparability of the target population. Comparable analysis samples can be achieved across the high school transcript studies by limiting analysis samples to high school graduates who received regular/standard or honors diplomas and

imposing additional restrictions such as earned credit minimums.

HS&B drew a national probability sample of high schools, as well as the sophomores and seniors within those schools, as of the 1980 spring term. By 1982, the school sample was no longer nationally representative (in the strictest sense) because it did not take into account school openings and closings in the 2-year period.

Similarly, while the HS&B senior cohort sample in 1980 generalized to the nation's high school seniors, the sophomore cohort in 1982 cannot be said to strictly represent the high school class of 1982. The HS&B sample was never freshened to add 1982 seniors who had no chance of selection 2 years before. This means that there is a bias in the HS&B 1982 (sophomores 2 years later) sample when it is used to generalize either to 12th-graders or to high school graduates. Seniors who were outside the United States 2 years before or seniors who were not sophomores 2 years before (e.g., seniors who repeated a year or who had a significantly accelerated trajectory) had no chance of selection into the sophomore cohort sample and are not represented within it.

The next two NCES high school cohort longitudinal studies, NELS:88 and ELS:2002, instituted a sample freshening procedure so that they include a nationally representative sample of high school seniors.

*Sample inclusion and exclusion***.** A second issue concerns student sample inclusion and exclusion, especially with respect to students with disabilities and English language learners.

In HS&B, sample members were classified as ineligible if deemed by their schools unable to complete the HS&B assessment battery owing to disability or lack of proficiency in English. Unfortunately, excluded students and specific reasons for exclusion were not well documented. However, it seems clear that the ineligible students represent the more severely disabled and the least proficient non-English speakers.

While some students were excluded from NELS:88, they were well documented, and over time their eligibility status was revisited. In ELS:2002, no students were excluded, though for those who could not complete survey forms, only contextual data and transcripts were collected. Also, in ELS:2002, some students received testing accommodations (e.g., extra time to complete the test); these cases are specially flagged.

Limiting 12th-grade high school graduate samples to recipients of regular or honors diplomas and eliminating cases that lack English course credits or that reflect a special education diploma or certificate of attendance largely eliminates the problem of differences in the excluded student population across studies. However, there is the remaining issue of how to identify and study the transcripts of individuals who had mild disabilities and how to compare the results over time. These issues arise because the longitudinal studies sought disability information from multiple respondent populations at multiple points in time. In NELS:88, for example, parents, teachers, students, and school administrators were all used as sources of information related to disability status.

Although some disability information is collected from sophomores' teachers, the primary source of identification for sophomore cohort members with disabilities in ELS:2002 is the Individualized Educational Program (IEP) flag, based on information taken from the sampling records provided by the base-year school, which identifies students in the school with IEPs.

Methodology across studies. In addition to differences in target populations and inclusion criteria, there are other differences among NCES high school transcript studies in terms of methodology. First, there is some variation in the statistical procedures used across studies. Overall, this variation will be the source of small differences that should not disrupt trend analyses. For example, different methods were used for nonresponse adjustment of weights. In HS&B, weighting cells were constructed based upon the known characteristics of the sample units. ELS:2002 used propensity modeling rather than a weighting cell approach. In NELS:88, a mix of the two approaches is encountered (propensity at the school level, weighting cells at the student level). However, results of nonresponse adjustment tend to be highly correlated regardless of method. Therefore, these differences should not lead to greatly different estimates.

*Content across studies***.** As curriculum changes, new courses emerge while others fall by the wayside. Therefore, with every transcript study, there is a need to add courses to the CSSC. Additionally, SST has been revised twice to accommodate changes in the curriculum. From a classification standpoint, adding new subject areas (such as information processing and computer studies) and expanded course offerings (including more AP courses) presents less of a quandary than certain efforts to achieve curriculum integration through interdisciplinary courses—confining such offerings (e.g., history of mathematics,

philosophy of science, psychological anthropology) in one subject category does injustice to certain aspects of the course content, while counting such courses in multiple areas may magnify and distort their impact.

As a result of these changes, many transcript composite variables have also changed over time. For example, with initiatives to seamlessly integrate academic into vocational education, conceptualizations of track or program type have changed. Such differences may reduce ease and simplicity in trend analysis, but are unavoidable features of the need to confront a complex and changing reality. Also, HS&B did not use as refined a system of course classification as did later studies, which, for example, distinguished courses based on whether they were remedial, regular, or advanced. On the other hand, some new measures developed out of NELS:88, such as the "pipeline" variables, which measure course content level, can be "read into" the other studies, such as the HS&B transcript studies.

The major limitation of these changes is that there are few coursetaking variables that are directly comparable across studies. For example, only a handful of courses qualified as computer science in the HS&B study. As the number of computer science courses has expanded, any variable based on computer science is not truly comparable across studies because it does not capture the range of courses that have emerged over time. Along with the two revisions of the SST, these changes make direct comparisons among coursetaking variables in the different files difficult. To facilitate some comparisons, ELS:2002 provides six summary measures that have directly comparable variables in NELS:88 and that can be constructed in HS&B by using existing elements. These variables are based on the same CSSC codes.

Analysts interested in comparing coursetaking patterns need to examine the CSSC codes available in each study. The CSSC codes are the same across studies, thus facilitating direct comparisons. As noted earlier, the list has evolved and certain subject areas have changed accordingly. Users may want to construct measures in a variety of ways to ensure that their findings are robust with respect to different variable specifications. In addition, analysts should consider changes in subject areas over time when conducting time trend analyses and interpreting findings.

There are many other variables that are typically linkable to transcripts; however, their status for this purpose may sometimes be problematic. For example, in HS&B and NELS:88, race was self-reported and students were asked to mark only one race. In light of the 2000 decennial census and revised race-reporting guidelines issued by the Office of Management and Budget, a new race category was added at the time of ELS:2002. More importantly, ELS:2002 respondents were allowed to mark all applicable races, thus generating a further category—multiracial. Knowing if a respondent who self-identified as Black on the HS&B questionnaire would have self-identified only as Black on the ELS:2002 questionnaire is impossible. To this extent, coursetaking trends for Blacks will be more uncertain than if a consistent definition had been maintained.

Test scores are another set of variables typically linked with transcript data that are different across studies. The relationship between coursetaking and tested achievement is of interest to researchers, and exploring the relationship between curriculum and assessment results is an interesting area for time series analysis. The NCES transcript studies provide only limited scope for such explorations.

Between NLS:72 in 1972 and ELS:2002 in 2004, the only subject consistently tested was mathematics. A further complication with comparative use of assessment data is changes in the measurement scale. Selectively, where content similarities permit, this limitation has been overcome by test linkage, usually IRT-based or equipercentile equating. One could, for example, examine the relationship between coursetaking and gain in the first 2 years of high school, using the equated 1980, 1990, and 2002 mathematics scores, or one could examine the relationship between coursetaking and gain for the periods 1990–92 and 2002–04, since ELS:2002 has been put on the NELS:88 scale. One final option for use of assessment data is to examine change within an effect size metric.

Contact Information

For content information about the High School Transcript Studies conducted as part of the Longitudinal Studies Program, 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

Methodology and Evaluation Reports

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.

Bozick, R., Lyttle, T., Siegel, 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 (*NCES 2006-338). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

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.

Ingels, S.J., Dowd, K.L., Taylor, J.R., Bartot, V.H., Frankel, M.R., Pulliam, P.A., and Quinn, P. (1995). *National Education Longitudinal Study of 1988 Second Follow-Up: Transcript Component Data File User's Manual* (NCES 95-377). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Planty, M., Bozick, R., and Ingels, S.J. (2006). *Academic Pathways, Preparation, and Performance: A Descriptive Overview of the Transcripts from the High School Graduating Class of 2003-04* (NCES 2007-316). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Chapter 30: Quick Response Information System

NCES has established two survey systems to collect time-sensitive, issue-oriented data quickly and with minimal response burden. The Fast Response Survey System focuses on collecting data at the elementary and secondary school levels. The Postsecondary Education Quick Information System collects data at the postsecondary level. These systems, subsumed under the general title, Quick Response Information System, are used to meet the data needs of U.S. Department of Education analysts, planners, and decisionmakers when information cannot be obtained quickly through traditional National Center for Education Statistics (NCES) surveys.

TWO QUICK RESPONSE INFORMATION SYSTEMS FAST RESPONSE SYSTEMS:

- Fast Response Survey System (FRSS) – 100 surveys since 1975
- Postsecondary Education Quick Information System (PEQIS) – 16 surveys since 1991

Fast Response Survey System

Overview

The Fast Response Survey System (FRSS) was established in 1975 to collect issue-oriented data quickly and with minimum response burden. The FRSS, whose surveys collect and report data on key education issues, was designed to meet the data needs of U.S. Department of Education analysts, planners, and decisionmakers when information could not be collected quickly through NCES's large recurring surveys. Findings from FRSS surveys have been included in congressional reports, testimony to congressional subcommittees, NCES reports, and other Department of Education reports. The findings are also often used by state and local education officials. From 1975 to 1990, the FRSS collected data at all education levels. Since the Postsecondary Education Quick Information System (PEQIS) was established in 1991, FRSS surveys have been limited to elementary and secondary school issues. To date, some 100 surveys have been conducted under the FRSS. Topics have ranged from racial and ethnic classifications at the state and school levels to the availability and use of resources, such as advanced telecommunications and libraries. Additionally, data have been collected on education reform, violence and discipline problems, parental involvement, curriculum placement and arts education, nutrition education, teacher training and professional development, vocational education, children's readiness for school, and the perspectives of school district superintendents, principals, and teachers on safe, disciplined, and drug-free schools. Some surveys, like surveys on Internet access and on teacher preparation and qualifications, have been conducted more than once in the past.

Data from FRSS surveys are representative at the national level, drawing from a universe that is appropriate for each study. Since 1991, the FRSS has generally collected data from public and private elementary and secondary schools, elementary and secondary school teachers and principals, public and school libraries, and, less frequently, state education agencies and local education agencies. Prior to 1991, FRSS also collected data from postsecondary institutions.

Sample Design

Data collected through FRSS surveys are representative at the national level, drawing from a universe that is appropriate for each study.

The FRSS collects data from state education agencies and national samples from other education organizations and participants, including local education agencies, public and private elementary and secondary schools, , and elementary and secondary school teachers and principals. To ensure minimal burden on respondents, the surveys are generally limited to three pages of questions, with a response burden of about 30 minutes per respondent. Sample sizes are relatively small (usually about 1,200 to 1,500 respondents per survey, but occasionally larger) so that data collection can be completed quickly.

The sampling frame for FRSS surveys is typically the NCES Common Core of Data (CCD) public school (or agency) universe. (See chapter 2.) The following variables are usually used for stratification or sorting within primary strata: instructional level (elementary school, middle school, and high school [secondary/combined]); categories of enrollment size; locale (city, urban fringe, town, rural); geographical region (Northeast, Southeast, Central, West); and categories of poverty status (based on eligibility for free or reduced-price lunch). The allocation of the samples to the primary strata is intended to ensure that the sample sizes are large enough to permit analyses of the questionnaire for major subgroups.

Within primary strata, the sample sizes are frequently allocated to the substrata in rough proportion to the aggregate square root of the size of enrollment of schools in the substratum. The use of the square root of enrollment to determine the sample allocation is considered reasonably efficient for estimating school-level characteristics and quantitative measures correlated with enrollment.

For example, the sample of elementary and secondary/combined schools for *Educational Technology in U.S. Public Schools: Fall 2008* was selected from the 2005–06 CCD Public School Universe data file, the most up-to-date file available at the time the sample was drawn. The sampling frame included over 85,000 regular schools. Excluded from this sampling frame were schools with a high grade of prekindergarten or kindergarten and ungraded schools, along with special education, vocational, and alternative/other schools; schools outside the 50 states and the District of Columbia; and schools with zero or missing enrollment.

The public school sampling frame was stratified by level (elementary or secondary/combined), categories of enrollment size, and categories for percent of students eligible for free/reduced-price lunch. Schools in the frame were then sorted by locale and region to induce additional implicit stratification. A sample of 2,010 schools were selected for the sample, but 56 were found to be ineligible for the survey because they were closed, merged, or did not meet the eligibility requirements for inclusion (e.g., they were special education, vocational, or alternative schools). This left a total of 1,950 eligible schools in the sample.

FRSS survey samples are sometimes constructed from the NCES Private School Universe Survey (PSS). (See chapter 3.) The sample usually consists of regular private elementary, secondary, and combined schools, with a private school being defined as a school not in the public system that provides instruction for any of grades 1–12 (or comparable ungraded levels) where the instruction is not provided in a private home. The following variables may be used for stratification or sorting within primary strata: instructional level (elementary, secondary, and combined), affiliation (Catholic, other religious, and nonsectarian), school size, geographic region, locale, and percentage of Black, Hispanic, and other race/ethnicity students. Schools are generally selected from each primary stratum with probabilities proportional to the weight reflecting the school's probability of inclusion in the area sample.

Other sources may serve as sampling frames, depending on the needs of the survey. For example, for *Participation of Migrant Students in Title I Migrant Education Program (MEP) Summer-Term Projects*, the districts and other entities serving migrant students were selected from the U.S. Department of Education's 1995–96 Migrant Education Program Universe data file.

Some FRSS surveys use a two-stage sampling process. For example, the *Teachers' Use of Educational Technology in U.S. Public Schools: 2009* and the *Educational Technology in Public School Districts: Fall 2008* which were administered concurrently with the *Educational Technology in U.S. Public Schools: Fall 2008* had a two-stage sampling process. The schools were selected during the first stage. The second stage of sampling for the Teacher Survey involved obtaining lists of teachers from the selected schools. The second stage of sampling for the Public School District Survey identified the districts that contained at least one of the sampled schools using the 2005-06 CCD Local Education Agency file.

Before PEQIS was established, the FRSS was sometimes used to examine postsecondary issues. For example, the *College-Level Remedial Education in the Fall of 1989* targeted institutions of higher education that served freshmen and were accredited at the college

level by an association or agency recognized by the U.S. Secretary of Education. The sampling frame was the universe file of the Higher Education General Information System (HEGIS) Fall Enrollment and Compliance Report of Institutions of Higher Education of 1983–84. (Note that HEGIS has since been replaced by the Integrated Postsecondary Education Data System—IPEDS—see chapter 12.) The universe of colleges and universities was stratified by type of control, type of institution, and enrollment size. Within strata, schools were selected at uniform rates, but the sampling rates varied considerably from stratum to stratum.

Data Collection and Processing

Most FRSS surveys are self-administered questionnaires where respondents are offered the option of completing the survey by mail or via the Web, with telephone follow-up for survey nonresponse and data clarification. On rare occasion a few have been telephone surveys, including one that used random digit dialing techniques. FRSS questionnaires are pretested, and efforts are made to check for consistency in the interpretation of questions and to eliminate ambiguous items before fielding the survey. For example, for the *Educational Technology in Public School Districts: Fall 2008* survey, questionnaires and cover letters were mailed to the superintendent of each sampled school district in early August 2008. The letter introduced the study and requested that the questionnaire be completed by the person most knowledgeable about educational technology in the district. Respondents were offered the option of completing the survey by mail or via the Web. Telephone follow-up for survey nonresponse and data clarification was initiated in late August 2008 and completed in January 2009.

Data are keyed with 100 percent verification. To check the data for accuracy and consistency, questionnaire responses undergo both manual and machine editing. Cases with missing or inconsistent items are recontacted by telephone.

Westat has served as the contractor for all surveys.

Estimation

Weighting. The response data are weighted to produce national estimates. The weights are designed to adjust for the variable probabilities of selection and differential nonresponse. Out-of-scope units are deleted from the initial sample before weighting and analysis. In the case of two-stage sampling—for example, in the *Teachers' Use of Educational Technology in U.S. Public Schools: 2009* —the weights used to produce national estimates were designed to reflect the variable probabilities of selection of the sampled schools and teachers and were adjusted for differential unit (teacher sampling list and questionnaire) nonresponse.

Imputation. Because item nonresponse rates in FRSS surveys are typically very low, the use of imputation is limited. The missing data are imputed using a "hot-deck" approach to obtain a "donor" from which the imputed values are derived. Once a donor is found, it is used to derive the imputed values for the missing data. For categorical items, the imputed value is simply the corresponding value from the donor. For numerical items, an appropriate ratio (e.g., the proportion of instructional rooms with wireless internet connections) is calculated for the donor, and this ratio is applied to available data (e.g., reported number of instructional rooms) for the recipient to obtain the corresponding imputed value. All missing items for a recipient are imputed from the same donor.

For example, in the *Educational Technology in U.S. Public Schools: Fall 2008* survey, all questionnaire items with response rates of less than 100 percent were imputed using the hot-deck imputation method. Under the "hot-deck" approach, a "donor" school that matched selected characteristics of the school with missing data (the recipient school) was identified. This survey used instructional level, categories of enrollment size, region, categories for percent combined enrollment of Black, Hispanic, Asian/Pacific Islander, or American Indian/Alaska Native students, categories for percent of students in the school eligible for free or reduced-price lunch, district size, and district poverty level as the matching characteristics. In addition, relevant questionnaire items were used to form appropriate imputation groupings. Once a donor was found, it was used to obtain the imputed values for the school with missing data. For categorical items, the imputed value was simply the corresponding value from the donor school. For the numerical items, an appropriate ratio was calculated for the donor school, and this ratio was applied to available data for the recipient school to obtain the corresponding imputed value.

Sampling Error

FRSS estimates are based on the selected samples and, consequently, are subject to sampling variability. The standard error is a measure of the variability of estimates due to sampling. Jackknife replication is the method used to compute estimates of the standard errors.

Nonsampling Error

Coverage Error. FRSS surveys are subject to any coverage error present in the major NCES data files

that serve as their sampling frames. Many FRSS surveys use CCD surveys as the sampling frame.

There is a 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 FRSS survey administration. Since teacher coverage depends on teacher lists sent by the schools, teacher coverage is assumed to be good. (See chapter 2 for a description of the CCD; see relevant chapters for other NCES surveys that serve as sampling frames for FRSS surveys.)

Nonresponse Error. Unit response for most FRSS surveys is 90 percent or higher. (See table 23.) Item nonresponse for most items is less than 1 percent. The weights are adjusted for unit nonresponse.

Measurement Error. Errors may result from such problems as misrecording of responses; incorrect editing, coding, and data entry; different interpretations of definitions and the meaning of questions; memory effects; the timing of the survey; and the respondent's inability to report certain data due to deficiencies in a recordkeeping system. Several specific examples of possible measurement error come from *the Public School Survey on Education Reform* and the *Public School Teacher Survey on Education Reform*, conducted in 1996. Survey results should be interpreted carefully for the following reasons: (1) survey questions were designed to be inclusive of a wide variety of reform activities since all principals and teachers do not share the same concept of reform; (2) respondents may have overreported activities in which they believe they should have been engaged; and (3) the questionnaire was too brief to collect information that could assist in judging the accuracy of the respondents' reports.

Data Quality and Comparability

Some FRSS surveys, such as surveys on internet access and on teacher preparation and qualifications, are repeated so that results can be compared over time. For example, the FRSS conducted the *Survey on Advanced Telecommunications in U.S. Public Schools* in 1994, 1995, 1996, and 1997. More recently, *Internet Access in U.S. Public Schools and Classrooms* was administered in 1998, 1999, 2000, 2001, 2002, 2003, and 2005. In addition, the *Survey on Advanced Telecommunications in U.S. Private Schools* was administered in 1995 and 1998–99. Results from the 1997 *Principal/School Disciplinarian Survey on School Violence* can be compared with those from the 1991 *Principal Survey on Safe, Disciplined, and Drug-Free Schools*, although there are some sampling differences that should be taken into account. (The 1997 survey was restricted to regular elementary and secondary schools, whereas the 1991 survey also included 13 vocational education and alternative schools in the sample.) Another example is provided by *Technology-Based Distance Education Courses for Public Elementary and Secondary School Students*, which was administered in 2002–03 and 2004–05. Two types of comparisons are possible with these FRSS data. The first type involves comparisons of the cross-sectional estimates for the two or more time periods. Cross-sectional comparisons reflect the net change in a given characteristic across years, including any changes in the underlying population. However, the enrollment estimates for 2002–03 and 2004-05 are different due to extensive data quality control procedures in place during data collection for the 2004–05 survey. The second type of comparison provides longitudinal analysis of change between 2002–03 and 2004–05. The longitudinal analysis is based on data from both administrations of the distance education survey, with the districts that existed both in 2002–03 and 2004–05 included in the analysis.

Occasionally, an FRSS survey is fielded to provide data that can be compared with data from another NCES survey. For example, the 1996 *Survey on Family and School Partnerships in Public Schools, K–8* was designed to provide data that could be compared with parent data from the 1996 National Household Education Survey as well as with data from the Prospects Study, a congressionally mandated study of educational growth and opportunity from 1991 to 1994. Another example is the 2001 *Survey on High School Guidance Counseling*, which was designed to provide data that could be compared to data from the 1984 Administrator and Teacher Survey supplement to the High School and Beyond Longitudinal Study.

Contact Information

For content information about the FRSS project, contact:

Peter C. Tice
Phone: (202) 502-7497
E-mail: peter.tice@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

Table 24. Weighted unit response rates for recent FRSS surveys: Selected years, 1999–2010

Survey	List participation rate	Weighted 1st level response rate	Overall weighted response rate
Teachers' Use of Educational Technology in U.S. Public Schools, 2009	81	79	65
Educational Technology in Public School Districts, Fall 2008	†	90	90
Educational Technology in U.S. Public Schools, Fall 2008	†	79	79
After-School Programs in Public Elementary Schools, 2008	†	91	91
Alternative Schools and Programs for Public School Students At Risk of Educational Failure, 2007-08	†	96	96
Distance Education Courses for Public School Elementary and Secondary School Students: 2004–05	†	96	96
Foods and Physical Activity in Public Elementary Schools: 2005	†	91	91
Public School Principals' Perceptions of Their School Facilities: Fall 2005	†	91	91
Internet Access in U.S. Public Schools and Classrooms: Fall 2005	†	86	86
Internet Access in U.S. Public Schools and Classrooms: Fall 2003	†	92	92
Internet Access in U.S. Public Schools and Classrooms: Fall 2002	†	90	90
Dual Credit and Exam-Based Courses: 2003	†	92	92
Distance Education Courses for Public School Elementary and Secondary School Students: 2002–03	†	96	96
Effects of Energy Needs and Expenditures on U.S. Public Schools: 2001	†	84	84
Survey on High School Guidance Counseling: 2001	†	94	94
District Survey of Alternative Schools and Programs: 2001	†	97	97

See notes at end of table.

Table 24. Weighted unit response rates for recent FRSS surveys: Selected years, 1999–2010–Continued

Survey	List participation rate	Weighted 1st level response rate	Overall weighted response rate
Survey of Classes that Serve Children Prior to Kindergarten in Public Schools: 2000–01	†	94	94
Survey on Programs for Adults in Public Library Outlets: 2000	†	97	97
Survey on Professional Development and Training in U.S. Public Schools: 1999–2000	88	85	75

† Not applicable.

SOURCE: Carver, P.R., and Lewis, L. (2010). *Alternative Schools and Programs for Students At Risk of Educational Failure, 2007-08* (NCES 2010-026). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Chaney, B., and Lewis, L. (2007). *Public School Principals' Report on Their School Facilities: Fall 2005* (NCES 2007-007). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Gray, L., and Lewis, L. (2009). *Educational Technology in Public School Districts, Fall 2008* (NCES 2010-003). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Gray, L., Thomas, N., and Lewis, L. (2010). *Educational Technology in U.S. Public Schools, Fall 2008* (NCES 2010-034). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Gray, L., Thomas, N., and Lewis, L. (2010). *Teachers' Use of Educational Technology in U.S. Public Schools, 2009* (NCES 2010-040). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Kleiner, A., and Lewis, L. (2003). *Internet Access in U.S. Public Schools and Classrooms: 1994–2002* (NCES 2004-011). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Kleiner, B., Porch, R., and Farris, E. (2002). *Public Alternative Schools and Programs for Students at Risk of Education Failure: 2000–01* (NCES 2002-004). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Lewis, L., and Farris, E. (2002). *Programs for Adults in Public Library Outlets* (NCES 2003-010). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Parsad, B., and Jones, J. (2005). *Internet Access in U.S. Public Schools and Classrooms: 1994–2003* (NCES 2005-015). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Parsad, B., and Lewis, L. (2006). *Calories In, Calories Out: Food and Exercise in Public Elementary Schools, 2005* (NCES 2006-057). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Parsad, B., and Lewis, L. (2009). *After-School Programs in Public Elementary Schools* (NCES 2009-043). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Parsad, B., Lewis, L., and Farris, E. (2001). *Teacher Preparation and Professional Development: 2000* (NCES 2001-088). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Setzer, J.C., and Lewis, L. (2005). *Distance Education Courses for Public Elementary and Secondary School Students: 2002–03* (NCES 2005-010). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Smith, T., Kleiner, A., Parsad, B., and Farris, E. (2003). *Prekindergarten in U.S. Public Schools: 2000–2001* (NCES 2003-019). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Smith, T., Porch, R., Farris, E., and Fowler, W. (2003). *Effects of Energy Needs and Expenditures on U.S. Public Schools* (NCES 2003-018). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Waits, T., Setzer, J.C., and Lewis, L. (2005). *Dual Credit and Exam-Based Courses in U.S. Public High Schools: 2002–03* (NCES 2005-009). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Wells, J., and Lewis, L. (2006). *Internet Access in U.S. Public Schools and Classrooms: 1994–2005* (NCES 2007-020). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Zandberg, I., and Lewis L. (2008). *Technology-Based Distance Education Courses for Public Elementary and Secondary School Students: 2002–03 and 2004–05* (NCES 2008-008). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Methodology and Evaluation Reports

Methodology is discussed in the technical notes to survey reports. Some of these reports are listed below.

Alexander, D., Heaviside, S., and Farris, E. (1999). *Status of Education Reform in Public Elementary and Secondary Schools: Teachers' Perspectives* (NCES 1999-045). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Carey, N., Lewis, L., and Farris, E. (1998). *Parent Involvement in Children's Education: Efforts by Public Elementary Schools* (NCES 98-032). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Carpenter, J. (1992). *Public School District Survey on Safe, Disciplined, and Drug-Free Schools* (NCES 92-008). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Carver, P.R., and Lewis, L. (2010). *Alternative Schools and Programs for Students At Risk of Educational Failure, 2007-08* (NCES 2010-026). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Chaney, B., and Lewis, L. (2007). *Public School Principals' Report on Their School Facilities: Fall 2005* (NCES 2007-007). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Gray, L., and Lewis, L. (2009). *Educational Technology in Public School Districts, Fall 2008* (NCES 2010-003). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Gray, L., Thomas, N., and Lewis, L. (2010). *Educational Technology in U.S. Public Schools, Fall 2008* (NCES 2010-034). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Gray, L., Thomas, N., and Lewis, L. (2010). *Teachers' Use of Educational Technology in U.S. Public Schools, 2009* (NCES 2010-040). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

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

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

Kleiner, A., and Lewis, L. (2003). *Internet Access in U.S. Public Schools and Classrooms: 1994–2002* (NCES 2004-011). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Kleiner, B., Porch, R., and Farris, E. (2002). *Public Alternative Schools and Programs for Students at Risk of Education Failure: 2000–01* (NCES 2002–004). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Lewis, L., and Farris, E. (2002). *Programs for Adults in Public Library Outlets* (NCES 2003-010). National Center for Education Statistics, Institute of Education Sciences, 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.

Parsad, B., Alexander, D., Farris, E., and Hudson, L. (2003). *High School Guidance Counseling* (NCES 2003-015). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Parsad, B., and Jones, J. (2005). *Internet Access in U.S. Public Schools and Classrooms: 1994–2003* (NCES 2005-015). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Parsad, B., and Lewis, L. (2006). *Calories In, Calories Out: Food and Exercise in Public Elementary Schools, 2005* (NCES 2006-057). National Center for Education Statistics, Institute of Education

Sciences, U.S. Department of Education. Washington, DC.

Parsad, B., and Lewis, L. (2009). *After-School Programs in Public Elementary Schools* (NCES 2009-043). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Parsad, B., Lewis, L., and Farris, E. (2001). *Teacher Preparation and Professional Development: 2000* (NCES 2001-088). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Setzer, J.C., and Lewis, L. (2005). *Distance Education Courses for Public Elementary and Secondary School Students: 2002–03* (NCES 2005-010). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Smith, T., Kleiner, A., Parsad, B., and Farris, E. (2003). *Prekindergarten in U.S. Public Schools: 2000–2001* (NCES 2003-019). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Smith, T., Porch, R., Farris, E., and Fowler, W. (2003). *Effects of Energy Needs and Expenditures on U.S. Public Schools* (NCES 2003-018). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Waits, T., Setzer, J.C., and Lewis, L. (2005). *Dual Credit and Exam-Based Courses in U.S. Public High Schools: 2002–03* (NCES 2005-009). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Wells, J., and Lewis, L. (2006). *Internet Access in U.S. Public Schools and Classrooms: 1994–2005* (NCES 2007-020). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Zandberg, I., and Lewis L. (2008). *Technology-Based Distance Education Courses for Public Elementary and Secondary School Students: 2002–03 and 2004–05* (NCES 2008-008). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Postsecondary Education Quick Information System

Overview

The Postsecondary Education Quick Information System (PEQIS) was established in 1991 to quickly collect limited amounts of policy-relevant information from a nationally representative sample of postsecondary institutions. Policy analysts, program planners, and decisionmakers in postsecondary education frequently need data on emerging issues quickly. It is not always feasible for NCES to use its large, recurring surveys to provide such data quickly, due to the length of time required to implement large-scale data collection efforts. In addition to obtaining information on emerging issues quickly, PEQIS surveys are used to assess the feasibility of developing large-scale data collection efforts on a given topic or to supplement other NCES postsecondary surveys. Surveys are generally limited to three pages of questions, with a response burden of about 30 minutes per respondent. To date, 16 PEQIS surveys have been completed, covering such diverse issues as distance learning, precollegiate programs for disadvantaged students, remedial education, campus crime and security, finances, services for deaf and hard-of-hearing students, and the accommodation of disabled students.

Sample Design

PEQIS employs a standing sample (panel) of approximately 1,600 nationally representative postsecondary education institutions at the 2- and 4-year levels. The panel includes public and private colleges and universities that award associate's, bachelor's, master's, and doctoral degrees. PEQIS can also conduct surveys of state higher education agencies. Four panels have been recruited since PEQIS was established in 1991. The sampling frame for the first PEQIS panel, recruited in 1992, was the 1990–91 Integrated Postsecondary Education Data System (IPEDS) Institutional Characteristics (IC) file. (See chapter 12.) The sampling frame for the second PEQIS panel, recruited in 1996, was the 1995–96 IPEDS IC file. The PEQIS panel was reselected in 1996 to reflect changes in the postsecondary education universe since the 1992 panel was recruited. A modified Keyfitz approach was used to maximize overlap between the 1992 and 1996 panels; this resulted in 80 percent of the institutions in the 1996 panel overlapping with the 1992 panel. The sampling frame for the third PEQIS panel, recruited in 2002, was the 2000 IPEDS IC file. A modified Keyfitz approach was used to maximize the overlap between the 1996 and 2002 samples; 81 percent of the institutions overlapped between these

two panels. The sampling frame for the 2006 PEQIS panel was the 2005 IPEDS IC file. The modified Keyfitz approach used to maximize the overlap between the 2002 and 2006 panels resulted in 79 percent of the institutions overlapping between the two panels.

Institutions eligible for the PEQIS frames for the 1992 and 1996 panels included 2-year and 4-year (including graduate-level) postsecondary institutions and less-than-2-year institutions of higher education. In 2002 and 2006, institutions eligible for the PEQIS frames were 2-year and 4-year (including graduate-level) Title IV eligible, degree-granting postsecondary institutions. In 1992, the sampling frame covered the 50 states, the District of Columbia, and Puerto Rico. In 1996 and subsequent years, institutions in Puerto Rico were excluded. The sampling frame included 5,320 institutions in 1992; 5,350 institutions in 1996; 4,180 institutions in 2002; and 4,270 institutions in 2006.

The sampling frames for all four PEQIS panels were stratified by instructional level (4-year and 2-year institutions for all four panels plus less-than-2-year institutions for the 1992 and 1996 panels); control (public, private nonprofit, private for-profit); highest level of offering (doctor's/first professional, master's, bachelor's, less than bachelor's); and total enrollment. Within each of the strata, institutions were sorted by region (Northeast, Southeast, Central, West), whether the institution had a relatively high percentage of Black, Hispanic, and other race/ethnicity students; and, in 1992 and 1996 only, whether the institution had research expenditures exceeding $1 million. The 1992 sample of 1,670 institutions was allocated to the strata in proportion to the aggregate square root of *full-time-equivalent enrollment.* The 1996 sample of 1,670 institutions was allocated to the strata in proportion to the aggregate square root of *total enrollment*, as was the 2002 sample of 1,610 institutions and the 2006 sample of 1,630 institutions. For all four panels, institutions within a stratum were sampled with equal probabilities of selection.

During recruitment for the 1992 panel, 50 institutions were found to be ineligible for PEQIS, primarily because they had closed or offered just correspondence courses. The final unweighted response rate at the end of PEQIS panel recruitment in spring 1992 was 98 percent (1,580 of the 1,620 eligible institutions). The weighted response rate for panel recruitment (weighted by the base weight) was 96 percent.

The modified Keyfitz approach used in 1996 resulted in 80 percent of the institutions in the 1996 panel overlapping with the 1992 panel. Panel recruitment was conducted with the 340 institutions that were not part of the overlap sample. Twenty institutions were found to be ineligible for PEQIS. The final unweighted response rate for the institutions that were not part of the overlap sample was 98 percent. The final participation rate across all 1,670 institutions selected for the 1996 panel was about 100 percent. The weighted panel participation rate (weighted by the base weight) was about 100 percent.

The modified Keyfitz approach used in 2002 resulted in 81 percent of the institutions in the 2002 panel overlapping with the 1996 panel. Panel recruitment was conducted with the 300 institutions that were not part of the overlap sample. During panel recruitment, 6 institutions were found to be ineligible for PEQIS. The final unweighted response rate at the end of PEQIS panel recruitment with the institutions that were not part of the overlap sample was 97 percent. There were 1,600 eligible institutions in the entire 2002 panel, because 4 institutions in the overlap sample were determined to be ineligible for various reasons. The final unweighted participation rate across the institutions selected for the 2002 panel was 99 percent (1,590 participating institutions out of 1,600 eligible institutions). The weighted panel participation rate was also 99 percent.

The modified Keyfitz approach used in 2006 resulted in 79 percent of the institutions in the 2006 panel overlapping with the 2002 panel. Panel recruitment was conducted with the 340 institutions selected for the 2006 panel that were not part of the 2002 panel. During panel recruitment, some institutions were found to be ineligible for PEQIS because they had closed. The final unweighted response rate at the end of PEQIS panel recruitment with the institutions that were not part of the overlap sample was 86 percent (290 of the 340 eligible institutions). There were 1,620 eligible institutions in the entire 2006 panel. The final unweighted participation rate across the institutions selected for the 2006 panel was 97 percent (1,570 participating institutions out of 1,620 eligible institutions). The weighted panel participation rate was 93 percent.

Data Collection and Processing

Typically, PEQIS surveys are self-administered questionnaires with respondents offered the option of completing the survey by mail or via the Web, with telephone follow-up for survey nonresponse and data clarification. Surveys are limited to three pages of questions, with a response burden of about 30 minutes per respondent. The questionnaires are pretested, and efforts are made to check for consistency in the interpretation of questions and to eliminate ambiguous

items before fielding the survey to all institutions in the sample.

The questionnaires are sent to institutional survey coordinators who identify the appropriate respondents for the particular survey and forward questionnaires to them. Nonrespondents who have not returned the survey within a set period of time are followed up by telephone. Data are keyed with 100 percent verification. To check the data for accuracy and consistency, questionnaire responses undergo both manual and machine editing. Cases with missing or inconsistent items are recontacted by telephone.

For the *Distance Education at Degree-Granting Postsecondary Institutions: 2006-07* survey, questionnaires were mailed to the PEQIS coordinators at the 1,630 institutions in fall 2007. The coordinators were told that the survey was designed to be completed by the person at the institution most knowledgeable about its distance education programs. In addition, data were collected from one 4-year private for-profit institution that was added to the sample for this survey only because it is the largest provider of online distance education courses in the nation, bringing the total sample size for this survey to 1,630 institutions. Respondents had the option of completing the survey online. Telephone follow-up of nonrespondents was initiated 3 weeks after mailout; data collection and clarification were completed in March 2008. Of the institutions that completed the survey, 72 percent completed it online, 20 percent completed it by mail, 5 percent completed it by fax, and 4 percent completed it by telephone.

Westat has served as the contractor for all surveys.

Weighting

The response data are weighted to produce national estimates. The weights are designed to adjust for the variable probabilities of selection and differential nonresponse. For recent PEQIS surveys, the weighted number of eligible institutions represents the estimated universe of approximately 4,240 Title IV-eligible degree-granting institutions in the 50 states and the District of Columbia.

Imputation

Item nonresponse rates for PEQIS surveys are typically very low (between 0 and 2 percent). Imputation was only performed for two surveys released before 2004; however, data have been imputed for all missing questionnaire data released thereafter. For the *Distance Education at Degree-Granting Postsecondary Institutions: 2006-07* survey, missing data were imputed using a "hot-deck" approach to obtain a "donor" institution from which the imputed values were derived. Under the hot-deck approach, a donor institution that matched selected characteristics of the institution with missing data (the recipient institution) was identified. Once a donor was found, it was used to derive the imputed values for the institution with missing data. For categorical items, the imputed value was simply the corresponding value from the donor institution. For numerical items, the imputed value was calculated by taking the donor's response for that item (e.g., enrollment in dual enrollment programs) and dividing that number by the total number of students enrolled in the donor institution. This ratio was then multiplied by the total number of students enrolled in the recipient institution to provide an imputed value. All missing items for a given institution were imputed from the same donor whenever possible.

Sampling Error

Estimates are based on the selected samples and, consequently, are subject to sampling variability. The standard error is a measure of the variability of estimates due to sampling. Because the data from PEQIS surveys are collected using a complex sampling design, the variances of the estimates from the surveys (e.g., estimates of proportions) are typically different from what would be expected from data collected with a simple random sample. To generate accurate standard errors for the estimates, standard errors are computed using a technique known as jackknife replication. The standard errors were calculated using a computer program.

Nonsampling Error

Nonsampling error describes variations in the estimates that may be caused by population coverage limitations and data collection, processing, and reporting procedures. The sources of nonsampling errors are typically problems like unit and item nonresponse, differences in respondents' interpretations of the meaning of questions, response differences related to the particular time the survey was conducted, and mistakes made during data preparation. It is difficult to identify and estimate either the amount of nonsampling error or the bias caused by this error. To minimize the potential for nonsampling error, the *Distance Education at Degree-Granting Postsecondary Institutions: 2006-07* survey used a variety of procedures, including a pretest of the questionnaire with the individual at each postsecondary institution deemed to be the most knowledgeable about its distance education programs and courses. The pretest provided the opportunity to check for consistency in the interpretation of questions and definitions and to eliminate ambiguous items. The questionnaire and instructions were also extensively reviewed by NCES and the data requestor at the Office of Educational Technology. In addition, both manual editing and machine editing of the questionnaire

responses were conducted to check the data for accuracy and consistency. Cases with missing or inconsistent items were recontacted by telephone to resolve problems. Data were keyed with 100 percent verification for surveys received by mail, fax, or telephone.

Coverage Error. Because the sampling frames for PEQIS surveys are constructed from IPEDS data files, coverage error is believed to be minimal.

Nonresponse Error. Both unit nonresponse and item nonresponse are quite low in PEQIS surveys. For the 16 surveys completed thus far, weighted unit response has ranged from 87 to 97 percent (see table 25). Item nonresponse for most items in PEQIS surveys has been less than 1 percent. The weights are adjusted for unit nonresponse.

For the PEQIS *Dual Enrollment of High School Students at Postsecondary Institutions: 2002-03*, 23 institutions were determined to be ineligible for the panel. For the eligible institutions, an unweighted response rate of 92 percent (1,460 responding institutions divided by the 1,590 eligible institutions in the sample for this survey) was obtained. The weighted response rate for this survey was 93 percent. The unweighted overall response rate was 91 percent (99 percent panel participation rate multiplied by the 92 percent survey response rate). The weighted overall response rate was 92 percent (99 percent weighted panel participation rate multiplied by the 93 percent weighted survey response rate).

Measurement Error. This type of nonsampling error may result from different interpretations of survey definitions by respondents or from the institution's inability to report according to survey specifications due to deficiencies in its recordkeeping system. Some examples of measurement error in PEQIS surveys follow.

For the PEQIS *Distance Education at Degree-Granting Postsecondary Institutions: 2006-07*, approximately 20 institutions were determined to be ineligible for the panel. For the eligible institutions, an unweighted response rate of 90 percent (1,450 responding institutions divided by the 1,610 eligible institutions in the sample for this survey) was obtained. The weighted response rate for this survey was 87 percent.

The 1995 Survey on Remedial Education in Higher Education Institutions was conducted to provide current national estimates on the extent of remediation on college campuses. Institutions provided information about the remedial reading, writing, and mathematics courses they offered in fall 1995. Remedial courses were defined as courses designed for college students lacking the skills necessary to perform college-level work at the level required by the institution. Thus, what constituted remedial courses varied by institution. Respondents were asked to include any courses meeting the definition, regardless of name. Some institutions refer to remedial courses as "compensatory," "developmental," or "basic skills."

Table 25. Weighted unit response rates for recent PEQIS surveys: Selected years, 2000–10

Survey	Panel participation rate	Weighted 1st level response rate	Overall weighted response rate
Distance Education at Postsecondary Institutions, 2006-07	†	87	87
Educational Technology in Teacher Education Programs for Initial Licensure	†	95	95
Dual Enrollment Programs and Courses for High School Students	99	93	92
Distance Education at Postsecondary Education Institutions, 2000-01	99	94	93

† Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics. (2009). Public-Use Data Files and Documentation (PEQIS 16): Distance Education at Postsecondary Institutions, 2006-07 (NCES 2009-074). U.S. Department of Education, National Center for Education Statistics. (2008). Educational Technology in Teacher Education Programs for Initial Licensure (PEQIS 15): Public-Use Data Files and Documentation (NCES 2008-013). U.S. Department of Education, National Center for Education Statistics. (2009). Public Use Data Files and Documentation (PEQIS 14): Dual Enrollment Programs and Courses for High School Students (NCES 2009-045). U.S. Department of Education, National Center for Education Statistics. (2005). Distance Education at Higher Education Institutions: 2000-01 (PEQIS 13): Public-Use Data Files and Documentation (NCES 2005-118).

Data Comparability

While most PEQIS surveys are not designed specifically for comparison with other surveys, the data from some PEQIS surveys can be compared with data from other postsecondary surveys. There have been, however, four administrations of the PEQIS *Survey on Distance Education Courses Offered by Higher Education Institutions* and two administrations of the *Survey on Remedial Education in Higher Education Institutions*.

The 1998 *Survey on Students With Disabilities at Postsecondary Education Institutions* complements another NCES study on the self-reported preparation, participation, and outcomes of students with disabilities. The latter study is based on an analysis of four different NCES surveys, which were used to address enrollment in postsecondary education, access to postsecondary education, persistence to degree attainment, and early labor market outcomes and graduate school enrollment rates of college graduates with disabilities. See *Students With Disabilities in Postsecondary Education: A Profile of Preparation, Participation, and Outcomes* (Horn and Berktold 1998).

The four administrations of the *Survey on Distance Education Courses Offered by Higher Education Institutions*—conducted first in late 1995, then in 1998–99, in 2000–01, and again in 2006–07—were the first to collect nationally representative data about distance education course offerings in higher education institutions. The four surveys differed in their samples and question wording. The sample for the first distance education survey, conducted in 1995, consisted of 2-year and 4-year higher education institutions in the 50 states, the District of Columbia, and Puerto Rico. At the time, NCES defined higher education institutions as institutions accredited at the college level by an agency recognized by the Secretary of the U.S. Department of Education. Higher education institutions were a subset of all postsecondary institutions. The sample for the second distance education survey, conducted in winter 1998–99, consisted of 2-year and 4-year postsecondary institutions (both higher education and other postsecondary institutions) in the 50 states and the District of Columbia. The third survey, conducted in 2000–01, included 2-year and 4-year Title IV-eligible, degree-granting institutions in the 50 states and the District of Columbia. Furthermore, data from the 1995 and 2000–01 surveys were not imputed for item nonresponse; however, comparisons between the surveys are possible when using the subset of higher education institutions from the 1998–99 survey. The fourth survey, conducted in 2006–07, also included 2-year and 4-year Title IV-eligible, degree-granting institutions in the 50 states and the District of Columbia. While this survey covered many of the same topics covered in the previous surveys, the data are not comparable because the definition of distance education in the 2006–07 survey reflected two major changes: First, the definition no longer included a criterion for instructional delivery to off-campus or remote locations; second, the definition included correspondence courses and distance education courses that were designated by institutions as hybrid/blended online courses.

The 1995 and 2000 administrations of the *Survey on Remedial Education in Higher Education Institutions* were conducted to provide national estimates on the extent of remediation on college campuses. The results update the information collected in two earlier NCES surveys for academic years 1983–84 and 1989–90; because PEQIS was not yet in existence, these surveys were conducted under the FRSS. (See section 1 of this chapter.) Although the 1995 survey was not designed as a comparative study, the results can be compared with data from the 1993–94 IPEDS Institutional Characteristics Survey: PEQIS estimated that 78 percent of institutions offered at least one remedial course for freshmen in fall 1995, and IPEDS estimated that 79 percent of institutions offered remedial courses in academic year 1993–94. At the student level, results from the 1995 PEQIS survey can be compared with results from institutional surveys conducted by the American Council on Education as well as a study conducted by the Southern Regional Education Board. However, these studies asked about freshmen needing remediation rather than about freshmen enrolled in remedial courses.

The remedial education data from the 1989 and 1995 surveys are not comparable to the data from the 2000 survey because of a change in the way that NCES categorized postsecondary institutions (and because of the inclusion of institutions in Puerto Rico in the earlier surveys). The data for the 1989 and 1995 surveys represent 2-year and 4-year higher education institutions that enroll freshmen. At the time these surveys were conducted, NCES defined higher education institutions as institutions accredited at the college level by an agency recognized by the Secretary of the U.S. Department of Education. Higher education institutions were a subset of all postsecondary institutions. The data for the 2000 survey represent 2-year and 4-year Title IV-eligible, degree-granting institutions that enroll freshmen. This change was necessary because the Department of Education stopped making a distinction between higher education institutions and other postsecondary institutions eligible to participate in federal Title IV financial aid

programs; thus, NCES no longer categorized institutions as higher education institutions. In order to make comparisons between the 1995 and 2000 surveys, the data from the 1995 survey can be reanalyzed with the definition of eligible institutions changed to match the definition for the 2000 survey as closely as possible.

Remedial enrollment can also be examined using postsecondary transcripts collected from institutions during the National Longitudinal Study of the High School Class of 1972 and the High School and Beyond Longitudinal Study (See chapters 6 and 7), as well as from student reported data in National Postsecondary Student Aid Study (NPSAS) (see chapter 14). Institutional reports of remedial enrollment in all of these surveys are substantially higher than student self-reports collected in NPSAS.

Contact Information

For content information on PEQIS, contact:

Peter C. Tice
Phone: (202) 502-7497
E-mail: peter.tice@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

Methodology and Evaluation Reports

Methodology is discussed in the technical notes to survey reports. Some of these reports are listed below.

Greene, B. (2005). *Distance Education at Higher Education Institutions: 2000–01* (NCES 2005-118). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Horn, L., and Berktold, J. (1998). *Students With Disabilities in Postsecondary Education: A Profile of Preparation, Participation, and Outcomes* (NCES 1999-187). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Kleiner, B., and Lewis, L. (2005). *Dual Enrollment of High School Students at Postsecondary Institutions: 2002–03* (NCES 2005-008). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Kleiner, B., Thomas, N., and Lewis, L. (2007). *Educational Technology in Teacher Education Programs for Initial Licensure* (NCES 2008-040). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Lewis, L., Alexander, D., and Farris, E. (1998). *Distance Education in Higher Education Institutions* (NCES 98-062). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Lewis, L., and Farris, E. (1997). *Campus Crime and Security at Postsecondary Education Institutions* (NCES 97-402). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Lewis, L., and Farris, E. (1999). *An Institutional Perspective on Students With Disabilities in Postsecondary Education* (NCES 1999-046). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Lewis, L., Snow, K., Farris, E., and Levin, D. (2000). *Distance Education at Postsecondary Education Institutions: 1997–98* (NCES 2000-013). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Parsad, B., and Lewis, L. (2003). *Remedial Education at Degree-Granting Postsecondary Institutions in Fall 2000* (NCES 2004-010). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Parsad, B., and Lewis, L. (2008). *Distance Education at Degree-Granting Postsecondary Institutions: 2006–07* (NCES 2009-044). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Phelps, R., Parsad, B., Farris, E., and Hudson, L. (2001). *Features of Occupational Programs at the Secondary and Postsecondary Education Levels* (NCES 2001-018). National Center for Education Statistics, U.S. Department of Education. Washington, DC.