NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-Up Data File Documentation

NCES 2014-361

U.S. DEPARTMENT OF EDUCATION

This page left intentionally blank.

NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

High School Longitudinal Study of 2009 (HSLS:09) First Follow-Up Data File Documentation

October 2013

Steven J. Ingels
Daniel J. Pratt
Deborah R. Herget
Jill A. Dever
Laura Burns Fritch
Randolph Ottem
James E. Rogers
RTI International

Sami Kitmitto
Steve Leinwand
American Institutes for Research

Elise Christopher
Project Officer
National Center for Education Statistics

NCES 2014-361 U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Easton
Director

National Center for Education Statistics
Jack Buckley
Commissioner

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

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

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

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

NCES, Institute of Education Sciences
U.S. Department of Education
1990 K Street NW, Washington, DC 20006-5651

October 2013

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

The NCES Home Page is http://nces.ed.gov.
The NCES Publications and Products address is http://nces.ed.gov/pubsearch.

This report was prepared for the National Center for Education Statistics under Contract No. ED-04-CO-0036 with RTI International. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation
Ingels, S.J., Pratt, D.J. Herget, D.R., Dever, J.A., Fritch, L.B., Ottem, R., Rogers, J.E.,Kitmitto, S., and Leinwand, S. (NCES 2014-361). *High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-Up Data File Documentation*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

For ordering information on this report, write to
U.S. Department of Education
ED Pubs
P.O. Box 2207, Alexandria, VA 22304

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

Content Contact
Elise Christopher
(202) 502-7899
elise.christopher@ed.gov

Foreword

This manual has been produced to familiarize data users with the design, and the procedures followed for data collection and processing, in the base year and first follow-up of the High School Longitudinal Study of 2009 (HSLS:09), with emphasis on the first follow-up. It also provides the necessary documentation for use of the public-use data files and information that will be helpful to analysts in accessing and understanding the restricted-use files.

Chapter 1 serves as an introduction to HSLS:09. It includes an overview and history of the National Center for Education Statistics (NCES) program of longitudinal high school cohort studies, summarizes the HSLS:09 objectives, and supplies an overview of the base-year and longitudinal study design.

Chapter 2 describes the base-year data collection instruments, including both the development and content of questionnaires (in the first follow-up, the student, parent, counselor, and school administrator questionnaires). Chapter 2 also provides information on the development of the mathematics assessment in algebraic reasoning and the scoring procedures and psychometric characteristics of this assessment.

The sample design used in the base year and first follow-up is documented in chapter 3. Data collection methods and results—including schedules, training, procedures, and response rates—are presented in chapter 4.

Chapter 5 describes data preparation and processing, including the receipt control system, coding operations, machine editing, and data file preparation. Additionally, chapter 5 provides information on the data preparation, scaling, and psychometric characteristics of some of the scales used in the first follow-up student and administrator surveys.

Chapter 6 describes weighting, variance estimation, and unit nonresponse bias estimation, while chapter 7 examines item-level statistical issues such as item nonresponse bias, imputation, and disclosure risk analysis. Chapter 8 describes the contents of the data files, including the data structure and linkages to other databases.

The appendixes include, among other topics, a list of the Electronic Codebook variables in the order of their appearance; a facsimile of first follow-up questionnaires, including flow charts and item wording; supplementary documentation for sample selection, imputed variables, bias analysis, and design effects; documentation of composite (derived) variables; and a glossary of terms.

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

Acknowledgments

The authors wish to thank the many individuals who helped conduct the study in the school setting, particularly principals and their staff, and school and information technology coordinators (school personnel who acted as liaisons to the study). The authors would also like to thank the many individuals who completed the survey instruments—student assessments and questionnaires, and questionnaires for parents, teachers, principals, and counselors—without whom this study would not be possible. Finally, we wish to acknowledge the contribution of the HSLS:09 First Follow-up Technical Review panel to the design of the study and its instruments.

Contents

This page left intentionally blank.

List of Tables

This page left intentionally blank.

List of Figures

Chapter 1. Introduction

1.1 Overview of the Data File Documentation

This data file documentation provides guidance and documentation for users of data from the base year through first follow-up of the High School Longitudinal Study of 2009 (HSLS:09). HSLS:09 is sponsored by the National Center for Education Statistics (NCES) of the Institute of Education Sciences, U.S. Department of Education, with additional support from the National Science Foundation.

The base-year and first follow-up studies were conducted through a contract to RTI International,[1] a university-affiliated, nonprofit research organization in North Carolina, in collaboration with its subcontractors, the American Institutes for Research, Horizon Research, Windwalker, and Research Support Services. This data file documentation contains information about the purposes of the study, the survey instruments, the mathematics assessment, the sample design, and the data collection and data processing procedures. The manual provides guidance for understanding and using all components of the base-year and first follow-up studies—student questionnaire and mathematics assessment data; questionnaire data from parents; and questionnaire data from mathematics and science teachers (teachers were surveyed only in the base year), school administrators, and counselors. Although this documentation supplies a highly detailed description of the first follow-up procedures and products, by far the most comprehensive source of information about the base year remains the Base-Year Data File Documentation (DFD) (Ingels, Pratt et al. 2011 [NCES 2011-328]). Data users are likely to need both DFDs to understand the base-year through first follow-up data.

The HSLS:09 dataset has been produced in both public-use and restricted-use versions. The publicly released data files reflect alteration or suppression of some of the original data. Such edits were imposed to minimize the risk of disclosing the identity of responding schools and the individuals within them. Although the main focus of this documentation is the public-use files, it contains much information relevant to the restricted-use data as well.

Chapter 1 addresses three main topics. First, it supplies an overview of the NCES secondary longitudinal studies program, thus situating HSLS:09 in the context of the earlier NCES high school cohorts studied in the 1970s, 1980s, 1990s, and 2000s. Second, it introduces HSLS:09 by delineating its principal objectives. Third, it provides an overview of the study design. In subsequent chapters, these additional topics are addressed: instrumentation (chapter 2), sample design (chapter 3), data collection methods and results (chapter 4), data preparation and processing (chapter 5), weighting and estimation (chapter 6), item nonresponse and imputation

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

(chapter 7), and data file structure and contents (chapter 8). Appendixes provide additional information, including a hardcopy version of the questionnaires, technical detail concerning sample selection, codebooks, and a glossary of terms.

1.2 Historical Background

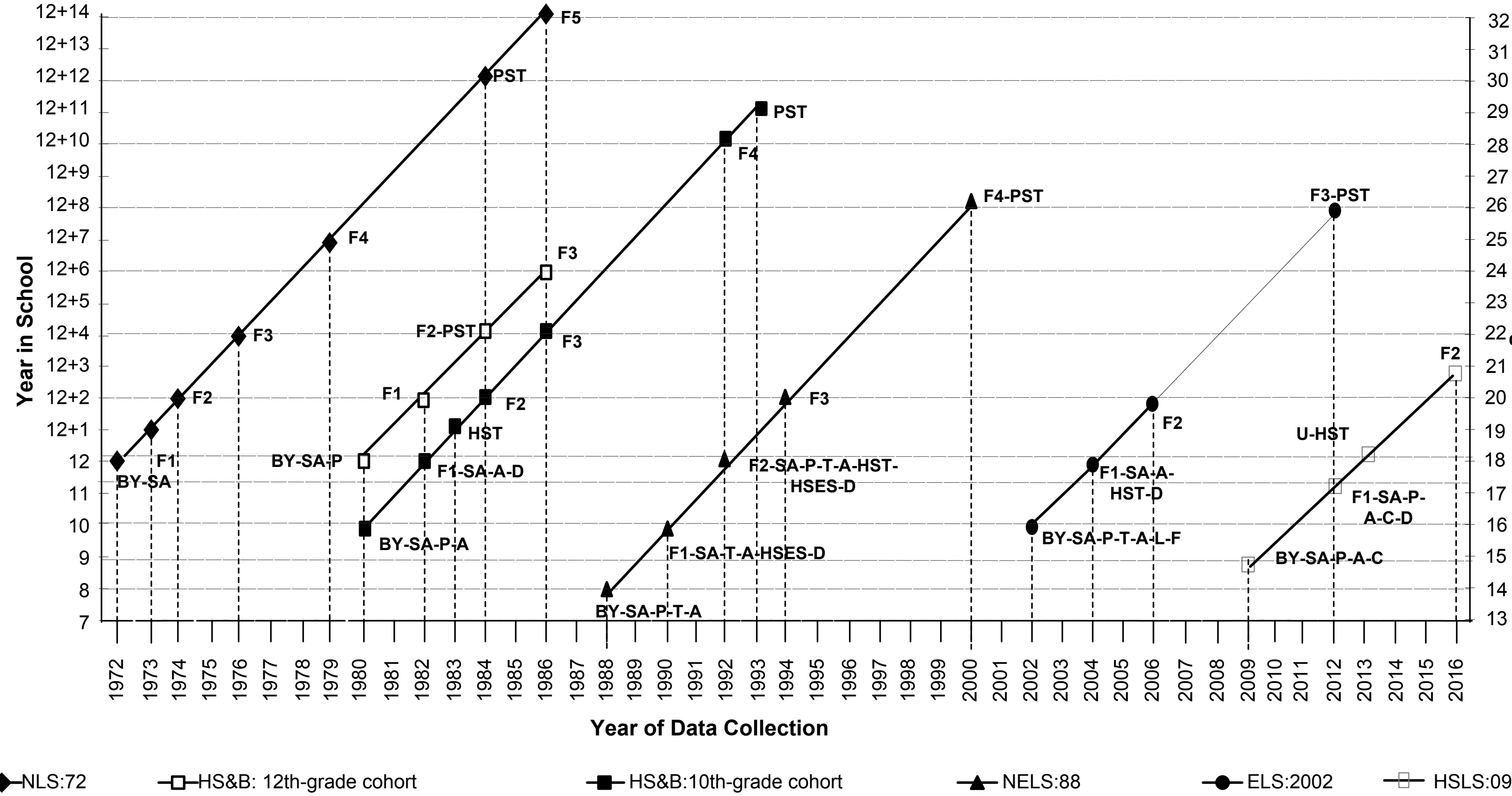
1.2.1 NCES Secondary Longitudinal Studies Program

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

The Secondary Longitudinal Studies program consists of three completed studies: the National Longitudinal Study of the High School Class of 1972 (NLS:72), the High School and Beyond (HS&B) longitudinal study of 1980, and the National Education Longitudinal Study of 1988 (NELS:88). In addition, base-year and first and second follow-up data for the Education Longitudinal Study of 2002 (ELS:2002)—the fourth longitudinal study in the series—are now available, and the ELS:2002 third follow-up (2012) survey data are in preparation for release. Taken together, these studies describe (or will describe) the educational experiences of students from four decades—the 1970s, 1980s, 1990s, and 2000s—and also provide bases for further understanding the correlates of educational success in the United States. These studies are now joined by a fifth longitudinal study—HSLS:09.

Figure 1 is a temporal representation of these five longitudinal education studies and highlights their component and comparison points for the time frame 1972–2016. (Currently, the next follow-up is scheduled to occur in 2016.)

Figure 1. Longitudinal design for the NCES high school cohorts: 1972–2016



NLS:72=National Longitudinal Study of the High School Class of 1972
HS&B=High School and Beyond: 1980
NELS:88=National Education Longitudinal Study of 1988
ELS:2002=Education Longitudinal Study of 2002
HSLS:09=High School Longitudinal Study of 2009

BY=Base-year data collection
F1=1st follow-up data collection
F2=2nd follow-up data collection
F3=3rd follow-up data collection
F4=4th follow-up data collection
F5=5th follow-up data collection

P=Parent survey
T=Teacher survey
A=Administrator survey
L=Library/media center survey
F=Facilities checklist
U=2013 update
HST=High school transcript
SA=Student assessment

PST=Postsecondary transcript
C=Counselor questionnaire
HSES=HS effectiveness study
D=Dropout survey

1.2.2 National Longitudinal Study of the High School Class of 1972

The Secondary Longitudinal Studies program began more than 40 years ago with the implementation of startup activities for NLS:72.[2] NLS:72 was designed to provide longitudinal data for educational policymakers and researchers to link educational experiences in high school with important downstream outcomes such as labor market experiences and postsecondary education enrollment and attainment. With a national probability sample of 19,001 high school seniors from 1,061 public and religious and other private schools, the NLS:72 sample was representative of approximately 3 million high school seniors enrolled in 17,000 U.S. high schools during the spring of the 1971–72 school year. Each student was asked to complete a student questionnaire and a cognitive test battery. In addition, administrators and counselors at the sample members' schools were asked to supply information about the schools' programs. Postsecondary education transcripts were collected in 1984 from the institutions attended by sample members. Five follow-up surveys were completed with this cohort, with the final data collection taking place in 1986, when the sample members were 14 years removed from high school and approximately 32 years old.

1.2.3 High School and Beyond

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

1.2.4 National Education Longitudinal Study of 1988

Much as NLS:72 captured a high school cohort of the 1970s and HS&B captured high school cohorts of the 1980s, NELS:88 was designed to study high school students of the 1990s—but with a baseline measure of their achievement and status, prior to their entry into high school.

---

[2] For documentation of the NLS:72 project, see Riccobono et al. (1981) and Tourangeau et al. (1987).

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

NELS:88 is an integrated system of data that tracked students from junior high or middle school through secondary and postsecondary education, labor market experiences, and marriage and family formation. Data collection for NELS:88 was initiated with the eighth-grade class of 1988 in the spring term of the 1987–88 school year. Along with a student survey, NELS:88 included surveys of parents (base year and second follow-up), teachers (base year, first and second follow-ups), and school administrators (base year, first and second follow-ups). The cohort was also surveyed twice after their scheduled high school graduation, in 1994 and 2000.[4] High school transcripts were collected in the fall of 1992 and postsecondary transcripts in the fall of 2000. Through a process of sample freshening, NELS:88 offers three nationally representative cohorts of students: spring-term 8th-, 10th-, and 12th-graders.

1.2.5 Education Longitudinal Study of 2002

ELS:2002 was designed to monitor the transition of a national sample of young people as they progress from 10th grade through high school and—as with its predecessor studies—on to postsecondary education or the world of work. ELS:2002 gathers information at multiple levels. In the base year (2002), it obtained information not only from students, but also from students' parents, teachers, and the administrators (principal and library media center director) of their schools. In the first follow-up (2004), the sample was freshened to represent the senior cohort of 2004 as well as the sophomore cohort of 2002, and high school transcripts were collected as were student questionnaires and tests and school administrator data.

In the second follow-up (2006), when most sample members had been out of high school for 2 years, computer-assisted student questionnaires were administered via the Web or telephone or in person, and data linkages and merges were added to the database, including SAT and ACT scores, General Educational Development scores, information from the Free Application for Federal Student Aid, and information from the National Student Loan Data System, including both federal loan and Pell grant information. Third follow-up questionnaire data (2012) are currently being processed, and postsecondary transcripts collected.[5]

[4] The entire compass of NELS:88, from its baseline through its final follow-up in 2000, is described in Curtin et al. (2002). NCES maintains an updated version of the NELS:88 bibliography on its website. The bibliography encompasses both project documentation and research articles, monographs, dissertations, and paper presentations employing NELS:88 data (see http://nces.ed.gov/surveys/nels88/Bibliography.asp).

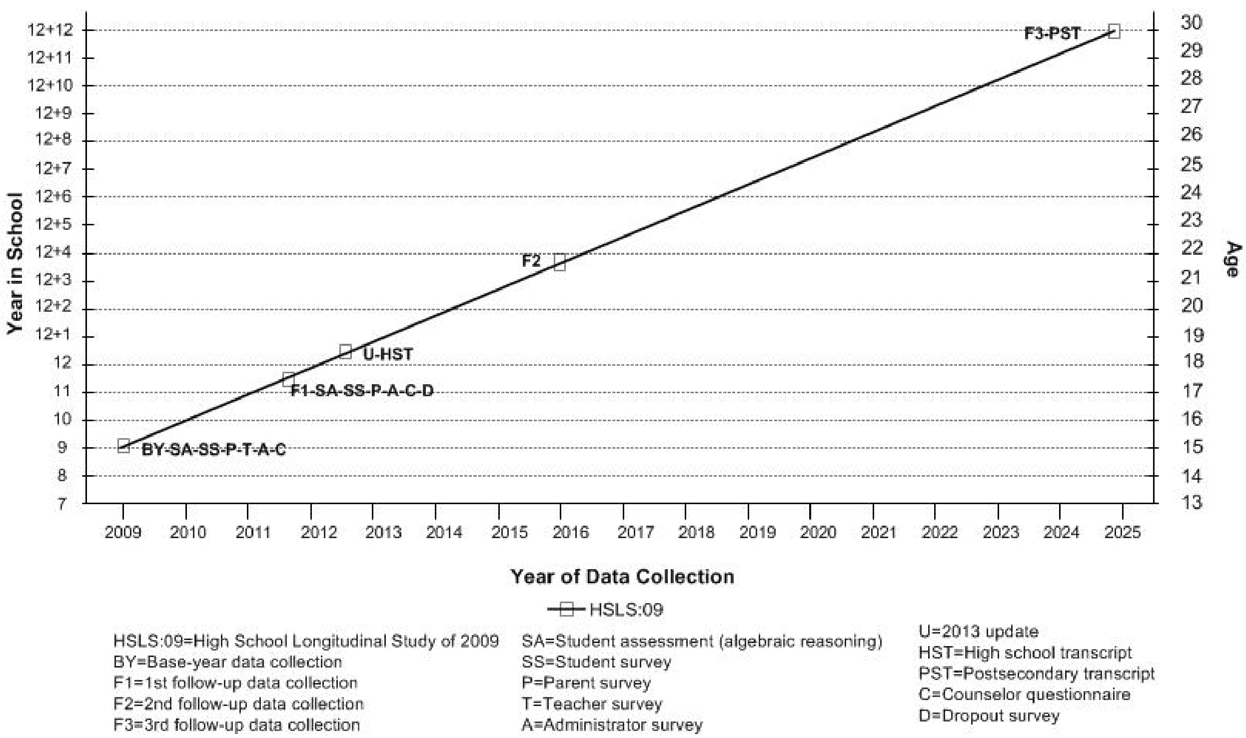
[5] ELS:2002 is documented in Ingels et al. (2007) (http://nces.ed.gov/pubsearch/pubsinfo.asp?/pubid=2008347). The third follow-up (2012) field test is documented in Ingels et al. (2012). A bibliography is maintained on the NCES website (http://nces.ed.gov/bibliography).

1.3 High School Longitudinal Study of 2009

1.3.1 Overview of the HSLS:09 Design and Objectives

The longitudinal design of HSLS:09 is set out in figure 2. The HSLS:09 base year took place in the 2009–10 school year, with a randomly selected sample of fall-term 9th-graders in more than 900 public and private high schools with both a 9th and an 11th grade.[6] Students took a mathematics assessment and survey online. Students' parents, principals, and mathematics and science teachers as well as the school's lead counselor completed surveys on the phone or on the Web.

Figure 2. Longitudinal design for the HSLS:09 9th-grade cohort: 2009–21



The first follow-up of HSLS:09 took place in 2012 when most sample members were in the spring term of the 11th grade. A postsecondary status update (2013 Update) is taking place in the summer and fall of 2013, to find out about the cohort's postsecondary plans and decisions. High school transcripts will be collected in the 2013–14 academic year, and a second follow-up is planned for 2016, when most sample members will be 3 years beyond high school graduation.

As these data points make clear, HSLS:09 departs from the design of the earlier cohorts in the secondary longitudinal studies series. HSLS:09 comprises a cohort of fall-term ninth-graders (in 2009). Its data collection points for students—fall term of 9th grade, spring term of

[6] Types of schools that were excluded from the sample based on the HSLS:09 eligibility definition are described as part of the discussion of the target population (see chapter 3).

(modal) 11th grade, 3 years out of high school—differ from the earlier collection points (spring of 8th grade, 10th grade, 12th grade; seniors 2 years out of high school) as well as in cohort definition. Very general comparison to the earlier studies (NLS:72, HS&B, NELS:88, and ELS:2002) is possible because all may be used to broadly model the transition from secondary schooling to the realm of postsecondary education and labor market participation and other adult roles. But the cross-cohort comparisons at specific grade levels that have been conducted in the past[7]—trends at the level of sophomores from 1980 to 2002; seniors from 1972 to 2004; sophomore cohort dropouts in 1982–2004; seniors 2 years later, 1974–2006—cannot be further extended with the HSLS:09 data. Nor does HSLS:09 employ the sort of sample freshening seen in NELS:88 and ELS:2002, through which multiple nationally representative grade cohorts were generated. HSLS:09 is, and is solely, a study of a fall ninth-grade cohort (including both ninth-graders who were repeating ninth grade and entering ninth-graders).

The core research questions for HSLS:09 explore secondary to postsecondary transition plans and the evolution of those plans; the paths into and out of science, technology, engineering, and mathematics (STEM); and the educational and social experiences that affect these shifts. (More will be said about research objectives in section 1.3.2 below.) The first follow-up data in particular are designed to facilitate the analysis of change, including gain in mathematical proficiency, and its correlates. Indeed—and unlike the base year—the first follow-up data cannot be used cross-sectionally because freshening for an 11th-grade cohort was not conducted. Given a 2009 ninth-grade cohort 2.5 years later, first follow-up data can only be used longitudinally.

HSLS:09 has both deep affinities and important differences with the prior studies, both of which will be highlighted in the discussion of study design below. Distinctive and innovative features of HSLS:09 include the following:

- starting point in the fall of 9th grade, the traditional beginning of high school;
- follow-up in spring of 11th grade, including follow-up math assessment;
- status update (through student or parent) summer-fall after the cohort's modal senior year;
- use of a computer-administered assessment and student questionnaire in a school setting;
- in first follow-up, questionnaire and assessment also computer-administered for out-of-school and transfer students;
- a two-stage multiform assessment that focuses on algebraic reasoning;
- use of computerized (web/computer-assisted telephone interview) parent (2009, 2012), teacher (2009), administrator (2009, 2012), and counselor (2009, 2012) questionnaires;

[7] See, for example, reports such as NCES 93-087, NCES 95-380, NCES 2006-327, NCES 2008-320, NCES 2009-307, and NCES 2012-345.

- inclusion of counselor surveys to document school course and program assignment policies and procedures, and counseling support of the transition to high school, and from high school to postsecondary education, training, and workforce participation;
- enhanced emphasis on the dynamics of educational and occupational decision-making;
- enhanced emphasis on STEM trajectories;
- augmentation of selected state public school samples to render them state-representative;

  Some differences with the predecessor secondary longitudinal studies should also be noted:

- unlike NELS:88 and ELS:2002, there is no freshening of students to create new grade cohorts, hence no cross-sectional level of analysis in the in-school follow-up round;
- unlike the prior studies, which tried to preserve content similarities to further trend analysis, HSLS:09 stresses new measures (as per topics noted above); and
- unlike the prior studies, which used, in large measure, the same fixed measurement points (e.g., spring term of senior year) HSLS:09 uses different measurement points (as noted in bullets above).

At the same time, there are also major points of continuity with all or several of the past studies:

- commitment to collecting high school (grades 9–12) transcripts as in HS&B, NELS:88, and ELS:2002 (although transcripts are for a ninth-grade class, not a high school graduating class, and **therefore are not usable for trend analysis**);
- a nationally representative school sample with an oversample of private schools and student numbers that are sufficient for subgroup reporting by major race/ethnicity categories, including Asians;
- commitment to following the cohort beyond high school;
- commitment to identifying and following high school dropouts;
- contextual samples of parents as in HS&B, NELS:88, and ELS:2002;
- contextual samples of teachers (in the base year) as in HS&B, NELS:88, and ELS:2002;
- a school administrator survey as in HS&B, NELS:88, and ELS:2002;
- an ability-adaptive assessment battery as in NELS:88 and ELS:2002; and
- production of a general purpose dataset that will support a broad range of descriptive and interpretive reporting.

1.3.2 HSLS:09 Research and Policy Issues

HSLS:09 addresses many of the same issues of transition from high school to postsecondary education and the labor force as were explored by its predecessor secondary longitudinal studies. At the same time, HSLS:09 brings a new and special emphasis to the study of youth transition by more intensely exploring the path that leads students to pursue and persist in courses and careers in the STEM fields.

With the advent of first follow-up data, HSLS:09 can now measure mathematics achievement gains in the first 3 years of high school. HSLS:09 can relate tested achievement to students' choice, access, and persistence—both in mathematics and science courses in high school, and thereafter in the STEM pipelines in postsecondary education and in STEM careers. Indeed, the HSLS:09 mathematics assessment serves not just as an outcome measure, but also as a predictor of readiness to proceed into STEM courses and careers. Because the study started with fall ninth-graders, it was able to identify high school dropouts in the first follow-up. The antecedent data from the base year will enable researchers to study the process of school disengagement, and will include relatively "early" dropout, those who left as early as spring of ninth grade.

The first follow-up will fill in a second data point to further depict the circumstances and implications for later outcomes of process data on student decision-making. Generally, across both the base year and follow-up, the study questions students on when, why, and how they make decisions about high school courses and postsecondary options, including what factors, from parental input to considerations of financial aid for postsecondary education, enter into these decisions. Questionnaires focus on factors that motivate students for STEM coursetaking and careers.

The transition into adulthood is of special interest to federal policy and programs. Adolescence is a time of psychological and physical changes. Attitudes, aspirations, and expectations are sensitive to the stimuli that adolescents experience, and environments influence the process of choosing among opportunities. Parents, educators, and policymakers all share the need to understand the effects that the presence or absence of good educational guidance from the school, in combination with that from the home, can have on the educational, occupational, and social success of youth.

These patterns of transition cover individual and institutional characteristics. At the individual level the study will look into educational attainment and personal development. In response to policy and scientific issues, data will also be provided on the demographic and background correlates of educational outcomes. At the institutional level, HSLS:09 contextual data for understanding student outcomes focus on school effectiveness issues, including promotion, retention, and curriculum content, structure, and sequencing. In turn, these factors may be associated with students' choice of, and assignment to, different mathematics and science courses and achievement in these two subject areas.

In sum, HSLS:09 data will allow researchers, educators, and policymakers to examine motivation, achievement, and persistence in STEM coursetaking and careers. More generally, HSLS:09 data drive analyses of changes in young people's lives and students' connections with communities, schools, teachers, families, parents, and friends along a number of dimensions, including the following:

- academic (especially in math), social, and interpersonal growth;
- transitions from high school to postsecondary education, and from school to work;
- students' choices about, access to, and persistence in math and science courses, majors, and STEM careers;
- the characteristics of high schools and postsecondary institutions and their impact on student outcomes;
- baccalaureate and sub-baccalaureate attainment;
- family formation, including marriage and family development, and how prior experiences in and out of school relate to these decisions, and how marital and parental status affect educational choice, persistence, and attainment; and
- the contexts of education, including how minority and at-risk status is associated with education and labor market outcomes.

1.3.3 HSLS:09 Analysis Files and Systems

HSLS:09 base-year through first follow-up data are available in two distinct applications: a restricted-use (NCES 2014-359) and a public-use (NCES 2014-358) electronic codebook housed on a DVD; and an online Education Data Application Tool for public-use data. Details of file structure and contents across these applications are supplied in chapter 8.

Chapter 2. Base-year Through First Follow-up Instrumentation

2.1 Instrument Development: Goals, Processes, Procedures

The High School Longitudinal Study of 2009 (HSLS:09) is intended to be a general-purpose dataset; that is, it is designed to serve multiple policy objectives. Policy issues studied through HSLS:09 include the identification of school attributes associated with mathematics achievement, college entry, and career choice; the influence that parents, teachers, and peers have on students' achievement and development; the factors associated with dropping out of (and returning to) the education system; and the transition of different groups (e.g., racial and ethnic, sex, and socioeconomic status groups) from high school to postsecondary institutions and the labor market, and especially into science, technology, engineering, and mathematics (STEM) curricula and careers. HSLS:09 inquires into students' values and goals, factors affecting risk and resiliency, the social capital available to sample members, the nature of student interests and decision-making, and students' curricular and extracurricular experiences. HSLS:09 also includes measures of school climate; each student's native language and language use; student and parental education expectations; attendance at school; course and program selection; college plans, preparation, and information-seeking behavior; interactions with teachers and peers; as well as parental resources and support. The HSLS:09 data elements are designed to support research that speaks to the underlying dynamics and education processes that influence student achievement and development over time. HSLS:09 is first and foremost a longitudinal study (indeed, cross-sectional analyses are possible only in the base year); hence, survey items were chosen for their usefulness in predicting or explaining future outcomes as measured in later survey waves, and many base-year measures have been repeated in the first follow-up.

Instrument design for HSLS:09 was guided by a theoretical framework or conceptual model. In the base year, the theoretical framework or conceptual model (figure 3 in the base-year Data File Documentation [DFD] [Ingels et al. 2010 (NCES 2011-328)]) served as the starting point for identifying constructs to be measured. The baseline selections were made with the first follow-up and subsequent rounds explicitly in mind; that is, they anticipated that first follow-up item selection would be guided by the need to align theoretically hypothesized antecedents with outcomes, and to employ the first follow-up both to measure post-baseline outcomes and to obtain further data to help predict future outcomes.

More specifically, because HSLS:09 follows its cohort over time, outcomes in early stages of the study (e.g., tested achievement in the first follow-up) may serve both as outcomes and as potential predictors for outcomes measured later in the study (e.g., smoothness of transition to the workforce, or postsecondary educational access and choice). Because HSLS:09

is a longitudinal study with family and school contexts measured twice (through the base-year and first follow-up parent questionnaires, and through the base-year and first follow-up administrator and counselor questionnaires), these constructs can be updated to be period-specific. The first follow-up questionnaires, then, comprise measures repeated from the base year to measure change in a base-year construct (e.g., educational expectations) or outcome measures (e.g., dropping out of high school) that can be related to base-year antecedents, augmented by further items that are specific to the first follow-up (e.g., transition to high school ceases to be an emphasis in the first follow-up, but transition plans for postsecondary education loom larger).

Instruments were developed, and revised, based on results from the base-year and first follow-up field tests, cognitive interviews, and feedback from the Technical Review Panel (TRP) and Office of Management and Budget (OMB). Hands-on testing of the programming logic for the questionnaires (and the computerized assessment) constituted the final step in developing a survey-ready instrument, after content approval from OMB. Specific items for the mathematics assessments were reviewed by a mathematics advisory panel of mathematicians and mathematics educators (see section 2.3.1.1 in the base-year DFD). Assessment items are not reviewed by OMB, nor were specific assessment items reviewed by the TRP. However, the larger assessment framework and goals and the assessment results (as seen in overall item statistics from the field test) were an integral element of the TRP deliberations.

The field testing of procedures, questionnaires, and assessments was an especially important step in the development of the main study surveys. Field test instruments were evaluated in a number of ways. For the questionnaires, field test analyses included evaluation of item nonresponse, examination of test-retest reliabilities, calculation of scale reliabilities, and examination of correlations between theoretically related measures. For the achievement test in mathematics, item parameters were estimated for both 9th and 11th grade in the base-year field test, and reestimated in the first follow-up field test. Both classical and Item Response Theory (IRT) techniques were employed to determine the most appropriate items for inclusion in the final forms of the two stages of the test. Psychometric analyses included various measures of item difficulty and discrimination, investigation of reliability and factor structure, and analysis of differential item functioning (DIF). The base-year field test report is available from the National Center for Education Statistics (NCES) (Ingels et al. 2010 [NCES 2010-01]) and the first follow-up field test report is included as appendixes L and M to this document.

2.2 Questionnaire Content in the First Follow-up

The four first follow-up questionnaires are described in the section below. Hardcopy specifications of the electronic first follow-up questionnaires appear later in this document, as appendix A. Simplified hardcopy versions (lacking routing logic) can be viewed on the NCES HSLS:09 website (http://nces.ed.gov/surveys/hsls09/questionnaires.asp). First follow-up completed case criteria appear in section 2.2.2 of this chapter. Assessment design and scores are reported in this chapter, with more information on the development process in the first follow-up

Field Test Report (appendixes M and N of this document). For information on the base-year questionnaires, please see the base-year DFD (Ingels et al. 2010 [NCES 2011-328]).

2.2.1 First Follow-up Questionnaires

The first follow-up questionnaires build on the base year, and attempt to elaborate, expand on, and capture evolving attitudes and plans. There are, therefore, many repeated measures, such as math self-identity, science self-identity, math utility, science utility, math efficacy, science efficacy, and so on. At the same time, the questionnaires also contain new or updated items that respond to changes since the fall of ninth grade (e.g., greater emphasis on postsecondary planning, supplanting items on transition into high school).

The contents of the four first follow-up questionnaires—student, parent, administrator, and counselor—are described immediately below. Questionnaire facsimiles (and flow charts) for the first follow-up appear as appendixes to the DFD. Certain items were deemed “critical” (i.e., of special importance to the study), and respondents who skipped such items were prompted, with a message noting the importance of the item and requesting that they provide an answer if at all possible. Also, the critical items were typically given a special role in defining a completed case. For example, a certain percentage or number of critical items constituted a part of the criteria for completeness, as is explained further below (section 2.2.2). For each questionnaire a list of critical items may be found in appendix N.

First Follow-up Student Questionnaire. The “student questionnaire” is the instrument that targets the fall 2009 ninth-grade cohort members in the spring term of the 2011–12 school year, regardless of their school enrollment status (i.e., whether they are students, dropouts, or early graduates). The questionnaire was designed with content appropriate for dropouts and early graduates, as well as students still enrolled in the base-year school, those who have left the base-year school for homeschooling, or those who have transferred to a new school. The first follow-up student questionnaire was a web survey. Topics explored in the questionnaire include (but are not limited to) the following: high school attendance, grade progression, and attainment; school experiences (including withdrawal from school); demographics and family background (including plans and preparations for the future, particularly post-high school); completion of admissions tests; influences on thinking and behavior; related peer behaviors, expectations, and aspirations; college choice and characteristics; knowledge of tuition; and whether they think they will qualify for financial aid and whether they would apply and if not, why not. The questionnaire also asks about high school coursetaking (more detailed information will be obtained from the high school transcript study), with a focus on math and science; feelings about math and science classes; math and science identity and utility; extracurricular programs; time spent on homework; and jobs and work for pay.

Some first follow-up participants were nonrespondents in the base year. Therefore, a number of questions were asked only of sample members for whom the information was missing

in the base year. These items pertain to critical classification variables such as language use, and parental education and occupation.

Parent Questionnaire. In the first follow-up, a random subsample of students' parents were administered the parent questionnaire (for further information on the subsample, see chapter 3). Data collection staff asked that the parent or guardian most familiar with the school situation and experience of the student sample member complete the parent questionnaire. As with the student questionnaire, there are questions that have been adapted to the situation of parents of dropouts as well as to parents of school attendees. The first follow-up parent questionnaire was fielded as both a web survey and a computer-assisted telephone or in-person survey. In the beginning of data collection, the survey was fielded with the web option. As the period of data collection elapsed, parents were also given computer-assisted options, with an interviewer over the telephone and at the end of the field period an in-person interviewer.

Parents were asked relationship to their child, how much of the time the cohort member lives with the parent respondent, if other parents reside in the household, parent respondent's current marital status, counts of household members by age, school enrollment status of the student sample member, negative events, prior educational experience including grades their child repeated, school suspensions, dropout episodes, number of times parent contacted the school, family activities, parent-child activities to prepare for the postsecondary transition, parent aspirations and expectations, ability to complete a bachelor's degree, ranking of importance of various college features, degree of parent and student input for postsecondary decisionmaking, affordability of college, means of getting financial aid information, expectations for qualification for financial aid, obstacles to applying for financial aid, savings for education, willingness to borrow, family educational and occupational background, employment status, income, demographic background, and languages spoken in the home.

Also (as in the student instrument), some questions (e.g., national origin) that were asked in the base year are only repeated in the first follow-up if base-year data are missing owing to unit or item nonresponse.

School Administrator Questionnaire. In the first follow-up, the school administrator questionnaire targets the base-year schools, 2.5 years later. In addition, an abbreviated version of the administrator survey was fielded to collect information from schools to which students in the study transferred. The school administrator questionnaire was fielded as a web survey.

The full administrator survey consists of four sections: (1) school characteristics; (2) programs, policies, and statistics of the school; (3) school staffing; and (4) opinions and background of the school principal. The school characteristics section contains questions about the school type (e.g., regular, charter, alternative); magnet and school choice programs; academic calendar; course scheduling; hours of instruction; and percentage of students who attend area or regional career and technical schools. The second section includes questions about enrollment; proportion of students who receive free or reduced-price lunch, who are English language

learners, and who receive special education services; enrollment and assignment policies; average daily attendance; absenteeism policies; programs to help students who are struggling academically; credit recovery programs; alternative and dropout prevention programs; activities to increase student interest and achievement in math and science; and years of coursework required for graduation. The third section asks questions about teachers within the school. Topics include numbers of teachers by full- and part-time status and by subject matter; teacher recruitment and retention; teacher absenteeism rates; and support for new math and science teachers. The fourth and final section includes questions on counseling goals and emphases; difficulty and methods of filling teaching vacancies for science and math; principal's perception of school problems; principal demographic characteristics; and principal's educational background and experience.

Because the first three sections contain factual questions about the school, these questions could be completed by the principal or another knowledgeable individual designated by the principal. However, the final section contains background and subjective questions, and the only appropriate respondent is the principal. Therefore, different login credentials were issued to school administrators and their designees such that school administrators were able to access the entire questionnaire, while designees were able to access only the first three parts. In an effort to reduce the burden of reporting detailed statistics, respondents were instructed that informed estimates were acceptable.

In addition to the full school administrator questionnaire, an abbreviated version was sent to schools to which students had transferred. The abbreviated version could be completed by a knowledgeable person in the school administrator's office by web, computer-assisted telephone interview (CATI), or paper-and-pencil interview (PAPI) instruments. The abbreviated version included a subset of questions. These include questions about the school type (e.g., regular, charter, alternative); hours of instruction; course scheduling; enrollment; proportion of students who receive free or reduced-price lunch, who are English language learners, and who receive special education services; postsecondary destinations of seniors; numbers of full- and part-time teachers, and math and science teachers; and years of service of the principal.

School Counselor Questionnaire. As in the base year, in the first follow-up the head or senior-most counselor at each base-year school was asked to complete the survey. The resulting counselor data are purely contextual, linked to the basic unit of analysis, the student sample member. The student, in turn, will have no first follow-up counselor data if she or he transferred to a new school, went into homeschooling, or attended a base-year school in which the counselor did not participate in 2012. The school counselor questionnaire was fielded as a web survey.

The counselor survey contained four sections: (1) counselor staffing and practices, (2) programs and support for students, (3) math and science placement, and (4) school reporting and statistics on students. The first section includes questions on number of full- and part-time

counselors, average caseload, method of assignment to students, breakdowns of percentage of time spent between delivering various services to students, and counselor duties and functions.

The second section contains questions such as programs and supports offered by the school, dual or concurrent enrollment offerings, summer enrichment, sources of credits beyond those offered directly by the school, attention given students in need of extra assistance, dropout prevention programs and services it offers, General Educational Development preparation, assistance with college entrance exams, assistance identifying and applying to colleges and universities, modes of assistance with college or university applications or financial aid and Free Application for Federal Student Aid preparation, programs and initiatives to ease the transition from high school to work, percentage of juniors and seniors taking advantage of various work preparation services, and school linkages with local employers.

The third section includes questions such as factors associated with mathematics and science course placement and sequencing, importance of various factors for advanced science and math placement, onsite and offsite calculus and physics, student participation and success in Advanced Placement and International Baccalaureate courses and exams, and average SAT and ACT scores of the school. The last section contains questions about the types of transition and outcomes data collected and analyzed by the school.

2.2.2 First Follow-up Criteria for Defining Completed Interviews

A completed case was defined as a respondent having answered a combination of a certain percentage of critical items, a certain number of total items, and, for students, the completion of a mathematics assessment. Because of the nature of the web survey, respondents had the ability to answer or skip any item; therefore, completeness of data varies across respondents. However, parents and administrators could potentially respond to an abbreviated version of the survey. The amount of information required for inclusion on the data file reflected a dual requirement—evidence of respondent seriousness in responding to the survey, and data of substantive value.

Student. A student survey was counted as complete if any of the following criteria were met:

- at least 50 items were answered and at least 50 percent of the critical items were answered (see Appendix N on critical items for the student survey); or
- at least 30 items were answered and the mathematics assessment was completed (see section 2.3.2.3 for a definition of a complete math assessment); or
- at least 50 items were answered and analyst review found sufficient data to consider the case complete (e.g., contained critical items or clusters of items that could be used to address a research question).

Parent. A parent survey was counted as complete if any of the following criteria were met:

- at least 40 items were answered and at least two critical items were answered for web/CATI/computer-assisted personal interview respondents; or
- at least two critical items were answered for PAPI respondents (see appendix N for a list of critical items for the parent survey).

Administrator. An administrator survey was counted as complete if at least five critical items were answered (see appendix N for a list of critical items for the administrator survey).

Counselor. A counselor survey was counted as complete if at least 25 items were answered, and at least two critical items were answered (see appendix N for a list of critical items for the counselor survey).

2.3 HSLS:09 Mathematics Assessment of Algebraic Reasoning

This section describes the development of the HSLS:09 mathematics assessment of algebraic reasoning, the scoring procedures, and the types of scores reported. The HSLS:09 assessment battery provides measures at two time points of student achievement in algebra for a cohort of grade 9 students: the first assessment (HSLS:09 base year) was administered in 2009 during the fall term of the 9th-grade year; and the second (HSLS:09 first follow-up) was administered in the spring of 2012 when most of the cohort were in the second semester of their 11th-grade school year.[8] These measures of students' algebraic skills and understandings can be related to student background variables and educational processes, for individuals and for population subgroups. Assessment data for HSLS:09 will be used to study factors that contribute to individual and subgroup differences in achievement. Combined with HSLS:09 base-year data, the HSLS:09 first follow-up provides a longitudinal look at student achievement 2.5 years later.[9] The first follow-up assessment was administered in two settings: in-school (as in the base year) and out-of-school in a self-administered web-based environment. Data collection in the two settings is described in chapter 4, while psychometric issues are addressed in this chapter.

2.3.1 Mathematics Assessment Development in the Base Year

Detailed accounts of assessment development in the base year may be found in the Base-year Field Test Report (Ingels et al. 2010 [NCES 2010-01]) and the Base-year DFD (Ingels, Pratt et al. 2011 [NCES 2011-328]). However, a concise summary is provided in this section.

[8] Although most students took the assessment in the spring term of 2012, the assessment was administered from January to October 2012. Out-of-school student data collection was conducted between February and October 2012, while in-school student data collection occurred between January and June 2012. The data file contains the date that each student took the assessment.

[9] For examples of the use of an IRT-based score (estimated number-correct, or probability of proficiency) to measure change within similarly designed NCES longitudinal studies (Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 and Education Longitudinal Study of 2002 [ELS:2002]), see Guarino et al. (2006) and Bozick and Ingels (2008). The two NCES reports illustrate both principal approaches to measuring achievement gain within a regression framework: use of gain scores as the dependent variable (Guarino et al. 2006) versus use of follow-up scores as a covariate (Bozick and Ingels 2008).

2.3.1.1 Test Design and Format

To measure student achievement in algebra, changes in this achievement over time, and how this achievement is linked to a range of individual, home, and school factors, a set of items aligned to an algebraic reasoning framework was developed, reviewed, and used to guide item development.

2.3.1.2 Algebraic Reasoning Framework

The item development process began with the development of a set of test and item specifications that described the importance of algebra and defined the domain of algebraic reasoning for the mathematics assessment at both grades 9 and 11. This task entailed designing an assessment of student understanding, and growth in understanding, of key algebraic knowledge and skills in algebra as a measure of mathematical preparation for the study of science, preparation for further study within the mathematical sciences and statistics, and preparation for the requisite skills and expectations of the workplace. Accordingly, the framework was designed to assess a cross-section of understandings representative of the major domains of algebra and the key processes of algebra.

The test and item specifications describe six domains of algebraic content and four algebraic processes:

- Algebraic Content Domains:
  - The language of algebra
  - Proportional relationships and change
  - Linear equations, inequalities, and functions
  - Nonlinear equations, inequalities, and functions
  - Systems of equations
  - Sequences and recursive relationships
- Algebraic Processes:
  - Demonstrating algebraic skills
  - Using representations of algebraic ideas
  - Performing algebraic reasoning
  - Solving algebraic problems

Each item was coded to one of the Algebraic Content Domains and one of the Algebraic Processes.

2.3.2 Mathematics Assessment Development in the First Follow-up

The base-year administration of the HSLS:09 mathematics assessment was a two-stage adaptive assessment, composed of 73 unique items.[10] Originally, the assessment was designed to be fielded twice, once in the 2009–10 school year and again in the first follow-up 2011–12 school year. However, after scoring the base-year assessment, it was decided to extend the assessment by developing additional higher difficulty items. Such items would guard against ceiling effects while more accurately measuring the full spectrum of algebraic knowledge and skills that is taught and learned in the first 3 years of high school. Therefore, in spring 2011, a field test was conducted to develop such additional higher difficulty items.

Table 1 shows the spring 2011 first follow-up field-test design that enabled this broadening of the item pool and updating of the item statistics to create the HSLS:09 2012 assessment of algebraic reasoning. This field test was administered to 473 students with a goal of developing 20 new items, the best performing of which could be added to the first follow-up main study assessment.

Table 1. Spring 2011 grade 11 field-test design

	Form A	Form B	Form C	Unique items
New items		20		20
Previously allocated grade 11 items	12	12	12	36
Grade 9 main study items		8		8
Total	40	40	40	64

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up Field Test.

Further information about assessment development in the HSLS:09 first follow-up field test can be found in appendix L and appendix M of this document. Topics discussed include purposes and item development, assessment design, test-taker demographics, timing, completion rates, motivation and item performance, and classical and IRT item statistics.

2.3.2.1 Main Study Two-stage Computer-delivered Implementation of the First Follow-up Assessment

As with the base year, the HSLS:09 first follow-up mathematics assessment was administered by computer, using a two-stage design wherein each student completed a Stage 1 “router test” and then a Stage 2 test designated as “low,” “moderate,” or “high” difficulty that was assigned on the basis of Stage 1 performance. The first follow-up assessment consisted of 73 unique items, with 23 serving as linking items to the base-year assessment, with any given student receiving 40 items. Table 2 shows this design:

[10] Although 73 unique items were administered, owing to performance problems for one item, only 72 were scored, and the base-year scale has a range of 0–72.

- Each student took a common 15-item Stage 1 router test that consisted of 11 base-year linking items and 4 items unique to the first follow-up.
- On the basis of Stage 1 performance, each student was routed to a low, moderate, or high Stage 2 test, each consisting of 25 items.
- Items on the Stage 2 tests included 5 items linking the moderate and high tests; 12 base-year linking items were part of both the low and moderate Stage 2 tests.[11]
- Students were aware that they were taking a 40-item test in two parts, a 15-item part and a 25-item part.

The computer-delivered design included an online scientific calculator and allowed students to skip and return to items within each stage and to identify items for review within each stage before submitting their answers as finished.

Table 2. HSLS:09 mathematics assessment first follow-up main study design

Number of items administered per student			Number of items at Stage 1				Number of items at Stage 2			
Stage 1	Stage 2	Total	Unique	Base-year items	Total	Stage 2 level	Unique	Across levels	Base-year items	Total
15	25	40	4	11	15	Low	13	0	12	25
						Moderate	8	5		25
						High	20		0	25

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Of the 23 base-year items, 11 were included on the Stage 1 router, and 12 were included on both the Stage 2 low and Stage 2 moderate levels. In addition to these items, 50 items were selected from the augmented field-test pool to comprise the first follow-up Stage 1 router and Stage 2 tests based on the following criteria:

- Items needed to represent a balance across the six content domains and the four algebraic processes.
- The average difficulty of the 15 items allocated to the Stage 1 router test and to each set of 25 items on the Stage 2 tests was preset as follows on the basis of the difficulty parameter of the IRT model (i.e., the b-parameter; for details of IRT, see Hambleton and Swaminathan [1985]) obtained using the updated field-test data:
  - Stage 1 router average difficulty = 1.6
  - Stage 2, low test average difficulty < –0.6
  - Stage 2, moderate test average difficulty = 1.6
  - Stage 2, high test average difficulty > 2.6

[11] There were no items common to the low and high Stage 2 tests and there were no base-year linking items in the high Stage 2 test.

Additionally, students were assigned to the three Stage 2 tests on the basis of their Stage 1 router performance so that, based on field-test results, approximately 25 percent of students would be routed to the high form, 50 percent to the moderate form, and 25 percent to the low form.

When the items in the base-year administration are combined with the items in the first follow-up administration, there are 123 unique items administered across both waves: 73 items were administered in both the base-year assessment and first follow-up assessments with 23 items common across assessments. However, not all items were used in scoring. In the HSLS:09 base year, one item was eliminated from scoring on the basis of very weak item statistics. In the HSLS:09 first follow-up, four items were similarly eliminated. The number of unique items scored across both waves was 118: 23 items common across waves, 72 items used for scoring in the base year (49 unique to the base year), and 69 items used for scoring in the first follow-up (46 unique to the first follow-up). Scoring procedures are described further below.

2.3.2.2 First Follow-up Second-stage Form Assignment

A total of 18,507 students had analyzable assessment data. Table 3 shows the breakdown by form.

Table 3. Number and percentage of HSLS:09 first follow-up mathematics assessment test-takers, by form: 2012

		Percent	
Category	Number	Unweighted	Weighted[1]
Total	18,507	100.0	100.0
Second-stage form			
Low	4,787	25.9	29.9
Moderate	9,406	50.8	51.0
High	4,314	23.3	19.1

[1] Weighted estimates were calculated with the student analytic weight (W2STUDENT).

NOTE: Table does not include 232 assessment observations that were discarded from the analysis sample as a result of attempting too few items or for pattern marking.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

2.3.2.3 First Follow-up Main Study Scoring Procedures

The assessment data were examined for possible indicators of lack of motivation to answer questions to the best of the student's ability. Examples of possible indicators are missing responses and pattern marking (e.g., all answers were "A" or "ABCDABCDABCD…"). As presented in table 3, 18,507 students had analyzable data. Of the 18,739 students who took the assessment, 232 (1.2 percent) test records were discarded from the analysis sample for the following reasons:

- A total of 48 records were deleted for attempting (i.e., selecting one of the four response options for) fewer than six items.
- A total of 184 records were deleted for pattern marking (171 cases for selecting the same answer options to more than 10 consecutive items on a stage, 13 cases for having the repeating “ABCDABCDABCD” pattern on a stage).

Classical item analyses were then conducted to provide information on item performance. The classical item statistics including p+ value, adjusted item-test biserial correlations, omit rate, distractor statistics, and DIF statistics were computed and reviewed. The p+ value for each of the items is presented in appendix B of this document.

The scores used to describe students’ performance on the mathematics assessment are based on IRT (Hambleton and Swaminathan 1985). The IRT model uses patterns of correct, incorrect, and omitted responses to obtain ability estimates that are comparable across the low-, moderate-, and high-difficulty test forms. One of the assumptions under the IRT model is unidimensionality of the test items. To verify that the items met that assumption, confirmatory factor analysis (CFA) was conducted based on each first follow-up test form.[12] The model fit indices obtained from the CFA analyses suggested that the items were unidimensional within each form.

Specifically, the IRT three-parameter logistic (3PL) model was used to calibrate the test items and estimate a student’s ability. The 3PL model is a mathematical model for estimating the probability that a person will respond correctly to an item. This probability is given as a function of one parameter characterizing the proficiency of a given student and three parameters characterizing the properties of a given item—the item’s difficulty, discriminating ability, and a guessing factor. The IRT model accounts for the three characteristics of each test question in estimating a student’s ability. The item parameters for each of the items are presented in appendix B. BILOG-MG (Zimowski et al. 2003) was used in carrying out item calibration and student ability estimation. During item calibration, separate ability priors based on performance on the router test were used for each of the three subpopulations taking the different second-stage tests (i.e., low, moderate, and high forms). The Bayesian estimation procedure was applied in estimating student proficiency. As part of this step, four items that had been administered were eliminated from scoring on the basis of very weak item statistics. Three of these items were from the Stage 1 test and one was unique to the medium Stage 2 test.

IRT scoring has several advantages over traditional raw number-correct scoring. First, IRT uses the overall response pattern of right and wrong answers to estimate ability and therefore can account for the guessing factor—a low-ability student guessing several difficult items correctly. Specifically, if answers on several easy items are wrong, a correct difficult item

[12] It would be ideal to conduct the CFA based on the pool of all 69 scored items. However, because of the test design of this study, many item pairs had no common observations and therefore their covariance could not be computed. The large number of missing covariance will cause unreliable results if the CFA were based on the pool of all 69 items.

is assumed, in effect, to have been guessed. Second, unlike in raw number-correct scoring, where omitted (skipped) responses are treated as incorrect answers, IRT procedures number-correct treat the omitted responses as not administered and use the pattern of responses to estimate the probability of correct responses for all test questions. Therefore, omitted items are less likely to cause distortion of scores as long as enough items have been answered right and wrong to establish a consistent pattern. Finally, IRT scoring makes it possible to compare scores obtained from test forms of different difficulty, such as HSLS:09 first follow-up.

The final step of the scoring procedure was to equate the IRT scores from HSLS:09 first follow-up to the scale of HSLS:09 base year so that scores may be compared longitudinally. The common items between the HSLS:09 base year and first follow-up (present in both Stages 1 and 2 of the first follow-up test) allowed for this equating to be possible. The tests were equated using the Stocking and Lord procedure (Stocking and Lord 1983). The procedure allowed the base-year thetas to remain unchanged while the first follow-up thetas were equated to the existing base-year scale.

2.3.2.4 Score Descriptions and Summary Statistics

Several types of scores are used in the HSLS:09 first follow-up to describe students' performance on algebra, all derived from the IRT model. Specifically, the IRT model uses information obtained from all students' response patterns of right and wrong answers as well as characteristics of the assessment items to compute a student ability estimate, theta. This theta (ability) estimate provides the basis for all other types of scores derived thereafter. On the data file, users will find the following scores:

- Theta (and the standard error of measurement of theta)
- Estimated number-correct
- Proficiency probability scores
- T-score (standardized theta score)
- Quintile scores

Details of the scores are described below. We begin with theta, since the theta estimate is the basis for all scores, proceed to criterion referenced scores (IRT estimated number-correct, which measure change in the aggregate, and probability of proficiency scores, which disaggregate gain, associating it with clusters of items that mark distinct hierarchical knowledge and skills), and proceed to the norm-referenced scores (standardized theta or T-score, quintile scores). The choice of the most appropriate score for analysis purposes should be driven by the context in which it is to be used. Table 4 presents the variable name, description, and summary statistics for the IRT estimated number-correct scores.

Table 4. Various types of scores from HSLS:09 first follow-up mathematics assessment, by variable: 2012

Variable	Description	Range[1]	Weighted mean[1]	Weighted standard deviation
X2TXMTH	HSLS:09 first follow-up mathematics theta score	−2.60–4.50	0.55	1.134
X2TXMSEM	HSLS:09 first follow-up mathematics theta standard error of measurement	0.18–0.76	0.34	0.083
X2TXMSCR	HSLS:09 first follow-up mathematics IRT-estimated number-correct on pool of base year and first follow-up items using first follow-up theta score (0–118)	25.0–115.1	64.4	19.00
X2X1TXMSCR	HSLS:09 first follow-up mathematics IRT-estimated number-correct on pool of base year and first follow-up items using base year theta score (0–118)	25.1–103.8	54.0	15.66
X2TXMTSCOR	HSLS:09 first follow-up mathematics standardized theta score (T-score)	22.2–84.9	50.0	10.00
X2TXMQUINT	HSLS:09 first follow-up mathematics quintile of theta scores	1.0–5.0	—	—

— Not available.

[1] Weighted estimates were calculated with the student analytic weight (W2STUDENT) except for X2X1TXMSCR which was calculated with the student longitudinal weight (W2W1STU).

NOTE: *n* = 20,594 for each variable (except X2X1TXMSCR), which includes 18,507 observations with scores from the assessment and 2,087 observations with imputed scores; for X2X1TXMSCR, *n* = 18,623 observations with a base-year theta.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

2.3.2.5 Theta (Ability) Estimate and Its Precision (Standard Error of Measurement of Theta)

Theta scores estimate ability in a particular domain. The theta scores are on the same metric as the IRT item-level difficulty parameters. Therefore, the theta scores may be less intuitively interpretable than a score that is a transformation of theta, such as the estimated number-correct or the T-score. However, the theta scores tend to be more normally distributed than estimated number-correct scores, because they are not dependent on the item difficulty parameters of the items within the scale score set. The standard error of measurement (SEM) of theta represents the precision of the IRT theta. The smaller the SEM, the greater the precision of measurement.

The theta (ability) scores provide a summary measure of achievement useful for correlational analysis with both status and educational process variables, such as demographics, school type, or behavioral measures (such as advanced mathematics coursetaking). They may be used in multivariate models as well, and provide measures of gain in algebraic reasoning ability over time.

2.3.2.6 IRT-estimated Number-correct Scores

The estimated number-correct score is an overall, criterion-referenced measure of achievement at a point in time. Based on the pattern of correct answers, the estimated number-correct score is an estimate of the number of items that students would have answered correctly had they responded to all items in the respective item pool defined for each criterion. However, the scores are not necessarily whole numbers, but typically include a decimal since they represent sums of probabilities.

The IRT-estimated number-correct scores are overall, criterion-referenced measures of status at a point in time. The criterion is the set of skills defined by the assessment framework and represented by the assessment item pool.

These scores may be used as longitudinal measures of overall growth, when an aggregated measure is preferred. (When a disaggregated measure is desired, to measure and compare gains made at different points on the score scale [i.e., to target a hierarchy of specific sets of skills], the probability of proficiency scores may be preferred in longitudinal analysis.) In the HSLS:09 first follow-up, the theta ability estimates and the item parameters derived from the IRT calibration were used to calculate each student's probability of a correct answer for each item in the pool. The sum of the probabilities across all 118 items is the estimate for the number of items the student would have gotten correct. Hence, the score has a potential range of 0 to 118.

Because the item pool increased because of additional new items being fielded in the first follow-up, the base-year number-correct scores were reestimated to reflect performance on the entire item pool of 118 unique items. A student's base-year theta, when available, was used to generate a base-year estimated number-correct while the first follow-up theta was used to generate a first follow-up estimated number-correct.

The IRT-estimated number-correct scores are useful in identifying cross-sectional differences among subgroups in overall achievement level (see Ingels, Dalton et al. 2011 [NCES 2011-327], table 4 for an illustration of the cross-sectional use of a variety of HSLS:09 mathematics scores). Similar to the theta (ability) scores above, they also provide a summary measure of achievement useful for correlational analysis with status and educational process variables, such as demographics, school type, or behavioral measures, and may be used in multivariate models as well. With the availability of assessment scores at two points in time, the IRT-estimated number-correct scores will be used in longitudinal analysis, as a measure of change in the aggregate. For example, in ELS:2002, at 10th grade, the average mathematics number-correct was 46.7, while the score in 12th grade was 51.2, a gain of 4.5 points on the 0–81 scale (Bozick and Ingels 2008).

2.3.2.7 Probability of Proficiency Scores

As noted above, the probability of proficiency scores provide a disaggregated measure of gain that helps to determine where on the vertical scale growth is taking place, either for various subgroups, or for various educational processes that may relate to change.

The mathematics proficiency probability scores are criterion-referenced and are based on clusters of items that mark seven levels on the mathematics scale developed in HSLS:09:

- **Level 1: Algebraic expressions.** Students able to answer questions such as these have an understanding of algebraic basics, including evaluating simple algebraic expressions and translating between verbal and symbolic representations of expressions.
- **Level 2: Multiplicative and proportional thinking.** Students able to answer questions such as these have an understanding of proportions and multiplicative situations and can solve proportional situation word problems, find the percent of a number, and identify equivalent algebraic expressions for multiplicative situations.
- **Level 3: Algebraic equivalents.** Students able to answer questions such as these have an understanding of algebraic equivalents and can link equivalent tabular and symbolic representations of linear equations, identify equivalent lines, and find the sum of variable expressions.
- **Level 4: Systems of equations.** Students able to answer questions such as these have an understanding of systems of linear equations and can solve such systems algebraically and graphically and characterize the lines (parallel, intersecting, collinear) represented by a system of linear equations.
- **Level 5: Linear functions.** Students able to answer questions such as these have an understanding of linear functions, can find and use slopes and intercepts of lines, and can use functional notation.
- **Level 6: Quadratic functions.** Students able to answer questions such as these have an understanding of quadratic functions and can solve quadratic equations and inequalities and understand the relationship between roots and the discriminant.
- **Level 7: Log and exponential functions (geometric sequences).** Students able to answer questions such as these have an understanding of exponential and log functions, including geometric sequences and can identify inverses of log and exponential functions and when geometric sequences converge.

The levels are hierarchical in the sense that mastery of a higher level typically implies proficiency at the lower levels. The HSLS:09 proficiency probabilities are IRT-derived estimates and are computed using IRT-estimated item parameters. The probability of proficiency for a given student at a given level is calculated as the probability of getting correct at least three of the four items in a given cluster marking a proficiency level based on the student's estimated ability. Each proficiency probability represents the likelihood that a student would pass a given hierarchical proficiency level. Although clusters of four items anchor each mastery level, the probability of proficiency is a continuous score that does not depend on a student answering the

actual items in each of the clusters but rather on the probability of a correct answer on these items given the overall pattern of response on the items completed.

Researchers should note that the proficiency probabilities are not intended to be conceptualized as subscales. Inasmuch as the clusters of items mark similar content, they are not representations of a subscale of that content. They are clusters of items around a common difficulty parameter. One of the assumptions of the IRT model is unidimensionality. Therefore, proficiency probabilities are simply meant to convey information as to the likelihood of mastery of these clusters of items along a unidimensional scale.

In the base-year assessment, five mastery or proficiency levels were identified. With the addition of more difficult items in the first follow-up assessment, two additional levels were identified. Thus five levels are calculated for the baseline and seven proficiency levels are calculated for its longitudinal follow-up (see table 5 for the first follow-up proficiency levels).

Table 5. HSLS:09 first follow-up algebra probability of proficiency scores, by variable: 2012

Variable	Description	Mathematical definition	Range[1]	Weighted mean[1]	Weighted standard deviation
X2TXMPROF1	HSLS:09 first follow-up proficiency probability score: Level 1	Algebraic expressions	0.01–1.00	0.92	0.191
X2TXMPROF2	HSLS:09 first follow-up proficiency probability score: Level 2	Multiplicative and proportional thinking	0.01–1.00	0.75	0.335
X2TXMPROF3	HSLS:09 first follow-up proficiency probability score: Level 3	Algebraic equivalents	0.02–1.00	0.64	0.372
X2TXMPROF4	HSLS:09 first follow-up proficiency probability score: Level 4	Systems of equations	0.03–1.00	0.29	0.335
X2TXMPROF5	HSLS:09 first follow-up proficiency probability score: Level 5	Linear functions	0.03–1.00	0.19	0.303
X2TXMPROF6	HSLS:09 first follow-up proficiency probability score: Level 6	Quadratic functions	0.02–1.00	0.05	0.119
X2TXMPROF7	HSLS:09 first follow-up proficiency probability score: Level 7	Log and exponential functions (geometric sequences)	0.00–0.92	0.02	0.061

[1] Weighted estimates were calculated with the student analytic weight (W2STUDENT).

NOTE: *n* = 20,594.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Probability of proficiency scores may be used in a number of ways.[13] They may be used to locate the achievement of HSLS:09 first follow-up sample members and subgroups at various behaviorally defined skill levels. The mean of a proficiency probability score aggregated over a subgroup of students is analogous to an estimate of the percentage of students in the subgroup who have displayed mastery of the particular skill. Because the range of the scores is 0 to 1, means can be expressed in percentage form. For example, the weighted mean for mastery of mathematics level 1 in the HSLS:09 first follow-up is 0.92, which is equivalent to saying that 92 percent of the first follow-up students had achieved mastery at this level (algebraic expressions). The proficiency probabilities are particularly appropriate for relating specific processes to changes that occur at different points along the score scale. For example, two groups may have similar gains, but for one group, gain may take place at an upper skill level, and for the other, at a lower skill level. To again use ELS:2002 as an example (see Bozick and Ingels 2008), aggregate 10th- to 12th-grade gains for blacks and whites were not statistically significantly different in the aggregate (a gain of 4.4 versus 4.5 points on the number-correct scale), yet for the proficiency levels, there was a difference in where the gains were taking place (for blacks, gains were at lower proficiency levels, for whites, gains were at higher proficiency levels—for example, despite similar gains in the aggregate, blacks gained more at levels 1 and 2 and whites gained more at level 4).

For those who gain at the higher skill level, there may be an association between their gains and curriculum exposure, such as taking advanced mathematics classes.

2.3.2.8 Standardized Scores (T-scores)

The standardized scores (T-scores) provide a norm-referenced measurement of achievement, that is, an estimate of achievement relative to the HSLS:09 first follow-up student population. They provide overall measures of status at a point in time compared with those of peers, as distinguished from the criterion-referenced scores, which represent status with respect to achievement on a particular criterion set of test items. The norm-referenced standardized scores do not answer the question "What skills do students have?" but rather, "How do they compare with their peers?"

The standardized T-score is a transformation of the IRT theta (ability) estimate, rescaled to a familiar metric with a mean of 50 and a standard deviation of 10. The transformation facilitates comparisons in standard deviation units. For example, an individual with a T-score of 75 (or a subgroup with a mean of 75) has performed 2.5 standard deviations above the national average for ninth-graders, whereas a score of 40 corresponds to 1 standard deviation below the

[13] See Ingels et al. 2011 (NCES 2011-328) for the cross-sectional use of probability of proficiency scores from the HSLS:09 base year. See Bozick and Ingels (2008) for an illustration of the use of probability proficiencies in a similar NCES longitudinal study, ELS:2002. For further discussion of the nonequivalence of scale score points and consequent need (if achievement gain is to be fully interpreted) for multiple criterion-referenced proficiency levels that mark distinct learning milestones, see Rock (2007, 2012).

norm. These numbers do not indicate whether students have mastered a particular algebraic skill or concept, but rather what their standing is relative to that of others. Further, because the scores are standardized within assessment, the base-year standardized T-score is not comparable to the first follow-up standardized T-score. The HSLS:09 first follow-up T-scores are documented in table 4, which also presents the summary statistics of the other types of scores discussed in this chapter.

2.3.2.9 Quintile Scores

The mathematics quintile score is a norm-referenced measure of achievement. The quintile score divides the weighted (population estimate) achievement distributions into five equal groups based on the mathematics standardized scores. Quintile 1 corresponds to the lowest achieving one-fifth of the population, quintile 5 the highest. To determine the quintile cut-points, the weighted distribution of the standardized scores was divided at the 20th, 40th, 60th, and 80th percentiles.

Quintile scores are convenient normative scores for the user who wants to focus on an analysis of background or process variables separately for students at different achievement levels. For example, one might want to compare the school experiences or educational aspirations of students in the lowest quintile with those of students in the highest quintile group. Table 4 contains the variable name, description, and range of values for the quintile scores.

2.3.2.10 Psychometric Properties of the Test

All items in the HSLS:09 first follow-up mathematics assessment pool were field tested. The field test was designed to provide information on item and test characteristics to ascertain the effectiveness of each item, develop a pool of main study items, and inform the placement of items on the main study test forms. Information about psychometric properties of the field-tested items, the setting of difficulty levels, DIF, and IRT scaling procedures are provided in the field-test reports (for the base year, Ingels et al. 2010 [NCES 2010-01]; for the first follow-up, appendixes M and N of this report). The validity of the content specifications for the HSLS:09 mathematics assessment was established by drawing from the NCTM standards, the mathematics frameworks of NAEP and TIMSS, and the review and approval of an expert panel.

The classical definition of reliability is the ratio of the true score variance to the observed score variance, which is the sum of the true scores variance and the error variance. In an IRT context, the true scores are the unobservable theta values that are estimated with a specified standard error from item response patterns. In the HSLS:09 first follow-up, where Bayesian estimation procedures were applied, the estimate of the error variance was computed as the mean of the variances of the posterior distributions of ability for each test-taker in the sample. The true score variance is estimated by the variance of the Bayesian theta scores (ability estimates) in the whole sample (see Bock and Mislevy [1982] for more information on Bayesian estimation). The reliability is therefore the true score variance divided by the sum of the true score variance and

the error variance (i.e., total variance). The IRT-estimated reliability of the HSLS:09 first follow-up test was 0.92 after sample weights were applied. This reliability is a function of the variance of repeated estimates of the IRT ability parameter (within variance), compared with the variability of the sample as a whole. This 0.92 reliability applies to all scale scores derived from the IRT estimation including the probability of proficiency scores. Imputed test scores were not included in the reliability estimation.

2.3.2.11 Comparison of Test Settings

One purpose of the field test was to determine whether it would be feasible to administer the mathematics assessment outside the school setting. Findings from the field test suggested that students (and perhaps dropouts) would make a genuine effort with the assessment in the out-of-school setting. For the main study assessment, of the 18,507 students whose assessment was scored, 15,236 (82.3 percent) were administered the assessment in a school setting and 3,271 (17.7 percent) were administered the assessment in an out-of-school setting. Table 6 provides characteristics of the students assessed in and out of school by enrollment status, race/ethnicity, and sex. By design, all students taking the assessment in school were enrolled, either in the same school that they were enrolled in the base year (99.7 percent) or (in the case of when several students had transferred en masse) in another school (0.3 percent). Similarly, all students who were home schooled, early graduates, or had left school by design had to take the assessment out of school. Students taking the assessment in school were more often White (57.8 percent) compared to those who took the assessment out of school (50.7 percent White).

Table 6. HSLS:09 first follow-up description of samples assessed, by assessment location: 2012

	All locations		In school (RTI laptop & Sojourn)		Out of school (web & CAPI)	
Characteristic	Number	Percent	Number	Percent	Number	Percent
All students	18,507	100.0	15,236	100.0	3,271	100.0
Enrollment status						
In base-year school	16,426	88.8	15,191	99.7	1,235	37.8
Transferred schools	1,553	8.4	45	0.3	1,508	46.1
Home school	145	0.8	N/A	N/A	145	4.4
Early graduate—regular completion	31	0.2	N/A	N/A	31	0.9
Early graduate—alternative completion	89	0.5	N/A	N/A	89	2.7
Left school	263	1.4	N/A	N/A	263	8.0
Race/ethnicity						
American Indian/Alaska Native, non-Hispanic	119	0.6	92	0.6	27	0.8
Asian, non-Hispanic	1,612	8.7	1,387	9.1	225	6.9
Black/African American, non-Hispanic	1,846	10.0	1,400	9.2	446	13.6
Hispanic, no race specified	288	1.6	218	1.4	70	2.1
Hispanic, race specified	2,571	13.9	2,039	13.4	532	16.3
More than one race, non-Hispanic	1,522	8.2	1,222	8.0	300	9.2
Native Hawaiian/Pacific Islander, non-Hispanic	77	0.4	66	0.4	11	0.3
White, non-Hispanic	10,468	56.6	8,810	57.8	1,658	50.7
Sex						
Male	9,266	50.1	7,687	50.5	1,579	48.3
Female	9,241	49.9	7,549	49.5	1,692	51.7

NOTE: n = 18,507; CAPI = computer-assisted personal interview. N/A = not applicable; counts and percentages in this table are not weighted.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 7 provides details on which second stage form students were assigned by assessment location. Students taking the assessment in school were more often routed to the high second stage form (24.2 percent) than students taking the assessment out of school (19.3 percent routed to the high form). Note that differences in routing to the second stage form on the assessment may be because of differences in the characteristics of those taking the assessment in each location type as noted above in table 6.

Table 7. HSLS:09 first follow-up second stage form, by assessment location: 2012

	All locations		In school (RTI laptop & Sojourn)		Out of school (web & CAPI)	
Characteristic	Number	Percent	Number	Percent	Number	Percent
All students	18,507	100.0	15,236	100.0	3,271	100.0
Second stage form						
Low	4787	25.9	3661	24.0	1126	34.4
Medium	9406	50.8	7893	51.8	1513	46.3
High	4314	23.3	3682	24.2	632	19.3

NOTE: *n* = 18,507; counts and percentages in this table are not weighted. CAPI = computer-assisted personal interview.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

The average number of items answered (out of 40 possible items on the assessment) is one measure of the amount of effort students put into the assessment, but may also be a result of students not reaching the end of the assessment. Students taking the assessment in school answered 38.9 items on average while students taking the assessment out of school answered 37.8 items on average (table 8).

Table 8. HSLS:09 first follow-up number of items answered, by assessment location: 2012

	All locations		In school (RTI laptop & Sojourn)		Out of school (web & CAPI)	
	Number	Mean	Number	Mean	Number	Mean
Number items answered	18,507	38.7	15,236	38.9	3,271	37.8

NOTE: *n* = 18,507; counts in this table are not weighted. CAPI = computer-assisted personal interview.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Finally, as noted previously, 232 student assessments had been discarded for possible lack of motivation. Of these 232 cases, 166 had been given the assessment in a school setting and 66 had been given the assessment in an out-of-school setting. Looking at these numbers in terms of percentages, overall 1.2 percent of cases were dropped for possible lack of motivation (232 of 18,739). For those given the assessment in a school setting, 1.1 percent were dropped (166 of 15,402) while for those given the assessment in an out-of-school setting, 2.0 percent were dropped (66 of 3,337).

Chapter 3. Sample Design

3.1 Overview of Sampling and Statistical Procedures

Details of the two-stage sample design for the High School Longitudinal Study of 2009 (HSLS:09) are provided below. The first follow-up sample consists of those students selected for the base year in 2009–10 that are still eligible for HSLS:09. Therefore, the base-year sample design is briefly reviewed first (section 3.2), followed by a detailed discussion of the changes introduced for the first follow-up design as it relates to the school and student samples (section 3.3). The sample design for selecting a random subsample of students' parents/guardians for the first follow-up concludes this chapter (section 3.3.4).

3.2 Base-year Sample Design

A summary of the complex design and resulting sample for the HSLS:09 base-year study is provided in this section. Section 3.2.1 contains a discussion of the stratified random selection of schools; section 3.2.2 documents the selection of students within schools; and section 3.2.3 describes the selection of the contextual samples. The base-year samples form the basis for the first follow-up samples discussed in section 3.3.

3.2.1 Selection of the School Sample

HSLS:09 is a stratified, two-stage random sample design with primary sampling units defined as schools selected in the first stage and students randomly selected from the sampled schools within the second stage. A total of 944 of 1,889 eligible schools participated in the base year resulting in a 55.5 percent weighted response rate (50.0 percent unweighted).

The HSLS:09 target population was defined in the base year as regular public schools, including public charter schools, and private schools in the 50 States and the District of Columbia providing instruction to students in both the 9th and 11th grades as of the fall of 2009. Excluded from this set of schools (i.e., study ineligible) were:

- Bureau of Indian Affairs schools;
- Special education schools for students with disabilities;
- Career technical education schools that do not enroll students directly;
- Department of Defense schools located outside the United States;
- Schools without both a 9th and 11th grade;
- Schools not in operation;
- Juvenile correction/detention facilities;

- Other schools that address disciplinary issues but do not enroll students directly;
- Ungraded schools (i.e., no metric to define students as being in the 9th grade);
- Schools that only offer testing services for home-schooled students; and
- Schools that do not require students to attend daily classes at their facility (e.g., online schools).

Initial base-year school samples were selected from two National Center for Education Statistics (NCES) files—public schools were selected from the 2005–06 Common Core of Data (CCD), and private schools were selected from the 2005–06 Private School Universe Survey (PSS). These files contained the most up-to-date information for sampling. Additional samples of new schools were selected just prior to recruitment from the updated 2006–07 CCD and 2007–08 PSS to maximize coverage of the target population.

A stratified probability proportional to size (PPS) sample of schools was selected from each sampling frame using a sequential probability with minimum replacement sampling algorithm (Chromy 1981). The school composite measure of size used in the sampling procedure was calculated as a linear combination of student counts multiplied by the desired overall sampling rates within race/ethnicity group (Folsom, Potter, and Williams 1987).

The 48 first-stage sampling strata were created by the interaction of school type (public, private–Catholic, private–other), geographic region (Northeast, Midwest, South, West), and geographic location of the school (i.e., metropolitan area or locale: city, suburban, town, rural).[14] Prior to drawing the random samples, the sampling frames were sorted by state within each of eight Census divisions (the nine Census divisions with two divisions collapsed into one) to ensure a representative distribution across the target population. The sample of 1,973 schools was allocated to the sampling strata in proportion to the relative number of ninth-grade students within the design strata.

HSLS:09 was originally designed to be representative of ninth-grade students in the 2009–10 school year in study-eligible schools across the United States (i.e., a national design). After receiving a request from the National Science Foundation for representative estimates within certain states, the design was augmented with additional sample schools to support the revised study objectives within 10 states. Results from a power analysis determined that at least 40 participating public schools per state would be sufficient to meet the precision criteria set for the national design. Additional schools for 8 of the 10 states were randomly selected through a Keyfitz (1951) procedure after the national sample was drawn to yield a total of 1,973 sample schools for the base-year study. However, during the recruitment phase of the study 84 schools were identified as ineligible, thus yielding 1,889 eligible schools.

[14] Locale is also referred to as urbanicity and is contained in the X1LOCALE variable discussed in the HSLS:09 base-year documentation.

3.2.2 Selection of the Student Sample

A sample of 26,305 students was randomly selected from the 944 participating schools in the base year. During base year recruitment, 1,099 students (4.2 percent unweighted) were classified as study ineligible and excluded from the data collection rosters, yielding 25,206 study-eligible students.

Students were randomly selected using a stratified systematic sampling procedure from base-year enrollment lists provided by administrative contacts at the school. The second-stage sampling strata were defined by the student's race/ethnicity (Hispanic, Asian, Black, and Other) specified by the school. The stratum-specific sampling rates were established to meet HSLS:09 analytic goals;[15] thus, Asian students were oversampled and an average of 28 ninth-graders were selected from each of the 944 participating schools.

Table 9 specifies the final response status for the base-year sampled students. A total of 21,444 students participated in the HSLS:09 base year (85.1 percent of the 25,206 eligible cases) by completing a questionnaire and mathematics ability assessment.[16] An additional 2.2 percent of the eligible students (or 548 cases) were classified by the school as being incapable of completing the student questionnaire because of physical limitations (e.g., visual impairment), cognitive disabilities, or limited English proficiency.

Table 9. Distribution of HSLS:09 study-eligible students, by base-year response status

Base-year response status	Count[1]	Percent
Study eligible	25,206	100.0
Study eligible, questionnaire capable	24,658	97.8
Respondent	21,444	85.1
Nonrespondent	3,214	12.8
Study eligible, questionnaire incapable	548	2.2
Physical limitations	38	0.2
Cognitive disabilities	303	1.2
Limited English proficiency	207	0.8

[1] A total sample of 26,305 students were originally selected from the 944 HSLS:09 participating schools. Study-ineligible students were identified after sampling (1,099 sample members) and excluded from further contact.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

[15] The HSLS:09 analytic goals included the calculation of efficient national estimates on characteristics associated with, for example, high school success and family influences in education choices. Sample size calculations determined that a minimum number of 21,000 students was required.

[16] Response in the base year was defined as those who completed a minimum number of sections in the student questionnaire (see the HSLS:09 Base-year Data File Documentation for details). As discussed later in chapter 6, a proportion of the responding students did not complete the math assessment. The mathematics ability (theta) and the associated standard error of measurement for the cases were imputed.

3.2.3 Selection of the Base-year Contextual Samples

Contextual information was collected on the student sample to describe the home and school environment. Home life and background information was obtained through students' parent questionnaires. School information was obtained through the students' administrator and counselor questionnaires. Students' teacher questionnaires (completed by science and mathematics teachers linked to the sampled student) captured information on teacher background and preparation, school climate, and subject-specific and classroom practices. The five contextual data sources for the HSLS:09 base year are discussed below. Summary information on the counts and response rates for each source of information including the contextual questionnaires is shown in table 10.

- *School Administrators*. Study recruiters contacted the administrators for all sampled HSLS:09 schools to request permission to conduct the in-school student data collection. Therefore, each school administrator was sampled with the same probability as calculated for the school.
- *School Counselors*. The lead counselor for the ninth-grade students at each of the 944 HSLS:09 schools participating in the base year was contacted to complete the counselor questionnaire. Therefore, as with the school administrator, the counselor was sampled with the same probability as calculated for the school.
- *Students' Parents and Guardians*. Contextual information on the student's home and family life was requested from one knowledgeable parent or guardian (referred to as "parent" in the subsequent discussion) for each of the 25,206 HSLS:09 study-eligible students. Therefore, the random selection probability for the parent was identical to that of his or her student.
- *Students' Science and Mathematics Teachers*. In addition to student and parent contact information, schools were requested to provide information on the science and math courses the student was currently attending. The associated math and science teacher for each sampled student (provided they were taking one or both of the courses) were contacted to complete a subject-specific questionnaire. Thus, teachers were randomly sampled with the same probability as the student.

Table 10. Summary of HSLS:09 base-year response rates: 2009

Instrument	Eligible	Participated[1]	Weighted percent[2]	Unweighted percent
Student questionnaire[3]	25,206	21,444	85.7	85.1
Student assessment[3]	25,206	20,781	83.0	82.4
Parent questionnaire	25,206	16,995	67.5	67.4
School administrator	25,206	23,800	94.5	94.4
School counselor	25,206	22,790	90.0	90.4
Teacher questionnaire[4]				
Mathematics teacher	23,621	17,882	71.9	75.7
Science teacher	22,597	16,269	70.2	72.0

[1] The participant counts exclude 548 questionnaire-incapable students.

[2] Percentages were calculated with the student base weight.

[3] Among questionnaire-capable students (n = 24,658), some 21,444 completed the student questionnaire, and 20,781 completed the mathematics assessment. Thus 87.0 percent (unweighted) completed the student interview or 87.4 percent weighted. Likewise, 84.3 percent (unweighted) completed a math assessment or 84.7 percent weighted.

[4] A total of 1,340 students did not have a science course and 471 did not have a mathematics course. In combination, 165 students had neither a science nor a mathematics course. These students were excluded from the calculations presented in this table.

NOTE: All percentages are based on the row under consideration.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

3.3 First Follow-up Sample Design

The first follow-up target populations are the same as defined for the base year. Consequently, the student target population contains all 9th-grade students as of fall 2009 who attended either regular public or private schools,[17] in the 50 United States and the District of Columbia, that provide instruction in both 9th and 11th grade. This population is referred to as the ninth-grade cohort in the subsequent discussions, where appropriate.

3.3.1 Base-year Schools and Transfer Schools in the First Follow-up

All of the 944 base-year participating schools were eligible for the HSLS:09 first follow-up. No new sample of schools was selected for this round. Therefore, the base-year school sample in the first follow-up is not representative of high schools with 9th and 11th grades in the 2011–12 school year, but rather is intended as an extension of the base-year student record, to be used to analyze school-level effects on longitudinal student outcomes.

Contact with the 944 sampled schools yielded 4 schools that were no longer in operation and 1 with no base-year sampled students. See table 11 for the distribution by school characteristics. Students associated with the study-ineligible schools are also included in the group of transfer students discussed below.

[17] Regular public schools also include public charter schools.

Table 11. Base-year weighted response rate and number of first follow-up eligible schools, by school characteristics: 2009–2012

School characteristics	Sample schools[1]	Number of base-year responding schools[1]	Base-year weighted response rate[2]	Number of first follow-up eligible schools[3]
Total	1,973	944	55.5	939
School type				
Public	1,550	767	58.8	765
Private (total)	423	177	46.2	174
Catholic	198	102	57.0	101
Other private	225	75	42.2	73
Region				
Northeast	357	149	40.9	148
Midwest	493	251	64.8	249
South	729	380	60.0	379
West	394	164	47.1	163
Locale				
City	667	272	44.1	270
Suburban	715	335	46.4	334
Town	204	117	67.5	116
Rural	387	220	66.6	219

[1] School characteristics are taken from the NCES files used for sampling. Namely, the 2005–06 Common Core of Data (CCD) and the 2005–06 Private School Universe Survey (PSS) for the initial sample of public and private schools respectively; and the subsequent sample of schools selected just prior to HSLS:09 base-year data collection from the 2006–07 CCD and the 2007–08 PSS.

[2] Weighted response rates were calculated with the school-level base weight (w_{1hi}) as the sum of the weights for the eligible, responding schools divided by the sum of the weights for all eligible schools in the HSLS:09 sample (see American Association for Public Opinion Research [2011] $RR1_w$).

[3] Counts include only those schools sampled in the base year and not any new study schools associated with transfer students.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

3.3.2 Student Sample

All 25,206 base-year study-eligible students, regardless of their response status, were included in the first follow-up sample. Unlike prior NCES studies, the HSLS:09 student sample was not freshened to include a representative later-grade cohort (such as 11th-graders in HSLS:09) as was done with 12th-graders in the Education Longitudinal Study of 2002.[18] Therefore, first follow-up estimates from the sample are associated only with the 9th-grade cohort 2.5 years later, and not the universe of students attending the 11th grade in the spring of 2012.

As with the school sample, some students originally labeled as eligible in the base year were found to be actually ineligible for HSLS:09. The distribution of questionnaire-capable

[18] See http://nces.ed.gov/surveys/els2002/.

students by sex and race/ethnicity is shown in table 12 for the complete base-year sample and the eligible samples within the base year and first follow-up.

Table 12. Base-year student response rates and number of first follow-up eligible students, by student characteristics: 2009–12

	Base-year study				
		Student participants			
Student characteristics	Eligible students	Number	Weighted percent[1]	Unweighted percent	First follow-up eligible students
Total	25,206	21,444	85.7	85.1	25,184
Sex					
Male	12,885	10,887	85.0	84.5	12,871
Female	12,321	10,557	86.4	85.7	12,313
Race/ethnicity					
American Indian/Alaska Native	249	223	87.1	89.6	249
Asian/Pacific Islander	2,576	2,144	86.2	83.2	2,575
Black or African American	3,115	2,684	86.8	86.2	3,111
Hispanic	3,958	3,516	88.6	88.8	3,954
White	14,702	12,630	86.2	85.9	14,690
Other race, more than one race, or missing value	606	247	34.4	40.8	605

[1] Weighted percentages use the student base weight.

NOTE: The variables used for sex and race/ethnicity are not presented on the main data file. To produce response rate calculations for all 25,206 eligible cases, information on sex and race/ethnicity relied on sampling frame variables that are not presented on the main data file.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

3.3.3 School-level Contextual Data

Four sources of school-level contextual information were collected for the students in the HSLS:09 base year—administrator, counselor, science teachers, and mathematics teachers (for those students enrolled in a science or mathematics course during the base year). Teacher responses were not collected in the first follow-up. The two school contextual sources remaining for this round of HSLS:09 (first follow-up administrator and counselor) are discussed below.

3.3.3.1 First Follow-up Administrator Survey

Questionnaires were sent to the administrator of the schools attended by any member of the HSLS:09 student sample. This included all of the 939 schools from the base year (base-year school) still in operation and still instructing at least one sampled student, and 1,822 schools not

included in the base-year sample where the sampled student transferred (transfer school). The distribution of the schools by these characteristics is discussed in chapter 4.[19]

3.3.3.2 First Follow-up Counselor Survey

Responses to the first follow-up counselor questionnaire were requested from a knowledgeable 11th-grade counselor at the 939 base-year schools. Transfer schools were excluded from this data collection component. Therefore, transfer student records will not contain any contextual data related to this instrument.

3.3.4 First follow-up Home-life (Parent) Contextual Data

Contextual information on the student's home background and family life was collected from one knowledgeable parent or guardian (hereafter referred to as "parent") in the base year for many HSLS:09 study-eligible students (section 3.2.3). The parent survey for the HSLS:09 first follow-up was administered to the parents of a random subsample of students as a cost-saving measure. This subsample was selected from all students who were eligible for the first follow-up, regardless of the student or parent base-year response status. The subsample design, including the distribution of the response status for the base-year parents, and associated sample sizes, is described below. Given that parents are brought into the sample through selection of students and represent a role rather than a single individual, it should be noted that the responding parent in the base year need not be identical to the parent in the first follow-up—for example, the respondent could be a student's mother in the base year, but a student's father could respond in the first follow-up, and so on. Once the students were subsampled, information was requested from the parent with the greater familiarity with the student's academic situation and home-learning environment. However, the responding adults were self-selected.

3.3.4.1 Subsample Design for the First Follow-up Parent Survey

Explaining changes in estimates from the base year to the first follow-up is of prime importance to researchers interested in HSLS:09. To ensure sufficient resources to maximize response from the sampled students, a decision was made to select a random subsample of parents in the first follow-up, with the goal of achieving 7,500 or more parent interviews.

The subsample of parents was randomly selected from within categories defined by the combination of the base year first- and second-stage sampling strata:

- school type (public, private-Catholic, private-other);
- 10 augmented-sample states (public schools only);
- the school's geographic region in the United States (Northeast, Midwest, South, West);

[19] Note that the transfer schools are not considered part of the school sample because they were included in the study only through a linkage with the HSLS:09 sample student. Additionally, administrator questionnaire responses are not available for responding students who no longer attended school in spring 2012 (i.e., early graduates, home-schooled students, dropouts).

- the school's locale in the United States (city, suburban, town, rural); and
- student race/ethnicity[20] (Hispanic, Asian, Black, Other).

A proportional allocation of the sample across the combined strata along with projected participation rates was used to determine the sampling rates within each group to yield a minimum of 7,500 completed parent questionnaires.[21,22] The proportional allocation ensured that the oversampling implemented in the base year (e.g., for Asian students) was preserved. Prior to selecting the sample, the student sampling frame records used in the base year were randomly sorted within state to ensure coverage across the characteristics of the base-year sample members not captured by the combined strata.[23]

The parent subsample was selected using a PPS minimal replacement methodology (Chromy 1981) and the student base weight as the measure of size. Use of the base weight from the base year minimized the variation in the first follow-up student home-life contextual base weights. This sampling approach has been used in other NCES surveys such as the National Education Longitudinal Study of 1988 fourth follow-up to subsample prior-wave nonrespondents (Liu and Aragon 2000).[24]

3.3.4.2 Sample Size for the First Follow-up Parent Survey

As discussed previously, *all* study-eligible students sampled in the base year were included in the first follow-up sample, and all 25,206 parents of base-year students were eligible for the parent survey subsample regardless of their base-year response status (table 13).

[20] The student race/ethnicity information was collected from the school enrollment lists and is available for all base-year sampled students.
[21] The subsample variables and proportions reflect student information known for the full sample because no information is available for the base-year nonresponding parents.
[22] Note that if the subsample was taken only from the set of base-year responding parents or base-year responding students, then any nonresponse bias existing in the initial estimates would also affect the first follow-up estimates. Subsampling from all base-year eligible sample members is one method to address this potential source of bias.
[23] For example, socioeconomic status (SES) is an important analytic variable used in many education surveys. SES, however, was only calculated for the HSLS:09 base-year participating students and therefore could not be used for the parent subsampling.
[24] See http://nces.ed.gov/surveys/nels88/.

Table 13. Base-year response status for HSLS:09 student-parent pairs

	Student response status						
	Respondent		Nonrespondent		Questionnaire incapable		
Parent response status	Number	Percent[1]	Number	Percent[1]	Number	Percent[1]	Total
Total	21,444		3,214		548		25,206
Respondent	16,429	65.2	403[2]	1.6	163	0.6	16,995
Nonrespondent	5,015	19.9	2,811	11.2	385	1.5	8,211

[1] Percentage of all records in each parent-student response status cell.

[2] Parent questionnaire data for nonresponding students were retained but not included on the public- or restricted-use files.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

Table 14 provides the distribution of the first follow-up parent survey subsample of 11,961 cases released for data collection. These counts include a primary parent survey subsample of 11,450 cases plus an additional 511 cases from the reserve subsample of 3,550.[25]

[25] A reserve sample of 3,550 parent cases were selected to ensure sufficient sample to obtain 7,500 completed cases. Cases within the combined strata were identified for release based on results from preliminary nonresponse bias analyses.

Table 14. Characteristics of the base-year parent sample compared with the first follow-up parent subsample

Characteristics	Base-year sample	First follow-up subsample	
		Number[1]	Percent[2]
Total	25,206	11,961	47.5
Parent base-year response status			
Respondent	16,592	8,086	48.7
Nonrespondent	8,614	3,875	45.0
School type			
Public	20,658	9,889	47.9
Private	4,548	2,072	45.6
Region			
Northeast	3,978	1,915	48.1
Midwest	6,673	3,047	45.7
South	10,207	4,819	47.2
West	4,348	2,180	50.1
Student base-year response status			
Respondent	21,444	10,216	47.6
Nonrespondent or questionnaire incapable	3,762	1,745	46.4
Student race/ethnicity[3]			
Hispanic	3,008	1,996	66.4
Asian	2,580	1,055	40.9
Black	2,890	1,341	46.4
Other/Unknown	16,728	7,569	45.2

[1] The first follow-up subsample includes an initial release of 11,450 cases plus an additional release of 511 cases from the reserve sample.

[2] Unweighted percent of base-year sample cases that were included in the first follow-up subsample.

[3] Student race/ethnicity is the four-category variable reported on the school rosters and used for base-year sampling. Updated race/ethnicity and other student demographic variables are only available for base-year respondents and therefore could not be used for parent subsampling.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Chapter 4. Data Collection Methodology and Results

4.1 Base-year Through First Follow-up Results Summary

Chapter 4 provides an account of data collection procedures implemented for the High School Longitudinal Study of 2009 (HSLS:09) first follow-up. (A detailed and comprehensive account of the base-year data collection may be found in the base-year Data File Documentation [DFD]: Ingels, Pratt et al. 2011, NCES 2011-328.) As a point of entry to this discussion, tables 15 and 16 summarize data collection results for each component in the two (2009 and 2012) survey waves.

Table 15. Summary of HSLS:09 base-year response rates: 2009

Questionnaire	Eligible	Participated	Weighted percent	Unweighted percent
School	1,889	944	55.5	50.0
School administrator[1]	944	888	94.9	94.1
School counselor[1]	944	852	91.3	90.3
Student questionnaire[2, 3]	25,206	21,444	85.7	85.1
Student assessment[2, 3]	25,206	20,781	83.0	82.4
Parent questionnaire[2]	25,206	16,995	67.5	67.4
School administrator[2]	25,206	23,800	94.5	94.4
School counselor[2]	25,206	22,790	90.0	90.4
Teacher questionnaire				
Math teacher[4]	23,621	17,882	71.9	75.7
Science teacher[5]	22,597	16,269	70.2	72.0

[1] Uses the school base weight.

[2] Uses the student base weight.

[3] Among questionnaire-capable students (n = 24,658), some 21,444 completed the student questionnaire, and 20,781 completed the mathematics assessment. Thus, 87.0 percent (unweighted) completed the student interview or 87.4 percent weighted. Likewise, 84.3 percent (unweighted) completed a math assessment or 84.7 percent weighted.

[4] Uses the student base weight. Results reflect students who were enrolled in a mathematics course.

[5] Uses the student base weight. Results reflect students who were enrolled in a science course.

NOTE: All percentages are based on the row under consideration.

SOURCE: U.S. Department of Education, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base-year.

Table 16. Summary of HSLS:09 first follow-up response rates: 2012

Questionnaire	Eligible	Participated	Weighted percent	Unweighted percent
Base-year schools[1]	939	904	†	96.3
School administrators	939	929	†	98.9
School counselors	939	925	†	98.5
Transfer school[2] administrators	1,822	1,346	†	71.5
Student questionnaire[3, 4]	25,184	20,594	82.0	81.8
Student assessment[3, 4]	25,184	18,507	73.0	73.5
Parent questionnaire[5, 6]	11,952	8,651	72.5	72.4
School administrator[4, 7]	23,432	22,498	95.4	96.0
School counselor[4, 7, 8]	20,858	20,601	98.6	98.8

† Not applicable. School sample is not representative of population.

[1] The HSLS:09 school sample included all schools that participated in the base-year data collection. However, five schools are not included in the number of eligible schools; four had closed and one did not have any base-year students still enrolled. Participating base-year schools include schools that conducted in-school student data collection sessions. All 939 schools were contacted to complete school administrator and school counselor questionnaires regardless of whether they conducted in-school student sessions.

[2] Transfer schools were identified from enrollment status updates provided by the school and responses provided in the student and parent questionnaire. Transfer schools were only contacted if at least one student from the transfer school participated in the first follow-up study.

[3] A total of 22 students from the base-year were ineligible for the first follow-up.

[4] Weighted percentage uses the student base weight.

[5] Weighted percentage uses the parent base weight.

[6] A subsample of 11,952 eligible parents were asked to participate in the HSLS:09 first follow-up data collection.

[7] The number eligible (23,432) refers to number of students linked to a school administrator. This number reflects all eligible students except for homeschooled students, as described in chapter 3. The number participated (22,498) refers to the number of eligible students linked to a school administrator who completed an administrator questionnaire.

[8] The number eligible (20,858) refers to number of students linked to a school counselor. This number reflects all eligible students except for homeschooled or transfer students, as described in chapter 3. The number participated (20,601) refers to the number of students linked to a school counselor who completed a counselor questionnaire.

NOTE: All percentages are based on the number of sample members in the row under consideration.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

4.2 First Follow-up Pre-data Collection Activities

All first follow-up instruments and procedures were fully pretested, and were revised on the basis of field test findings. A comprehensive account of the field test is found in the 2011 Field Test Report, which appears as appendixes L and M of this document.

The first follow-up data collection was scheduled to occur when most sampled students were in the spring of their 11th-grade year. Like the base-year data collection, contacting of school districts and schools began a year before data collection commenced. In-school data collection comprised a student questionnaire and mathematics assessment. Students who did not participate in the in-school session, including those who were no longer enrolled at the base-year

school, were contacted to complete the questionnaire and mathematics assessment outside of school. First follow-up data collection also included surveys of school administrators, counselors, and a subsample of parents. There was no teacher data collection in the first follow-up. Figure 3 shows the start and end dates of major HSLS:09 first follow-up activities.

Figure 3. Start and end dates for major HSLS:09 first follow-up activities: 2012

Activity	Start date	End date
School recruitment	January 2011	May 2012
Training of field staff	January 2012	January 2012
In-school student data collection[1]	January 2012	June 2012
Parent, school staff, and out-of-school student data collection	February 2012	October 2012

[1] Students who were contacted out of school were contacted through October 2012.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

4.2.1 Enlisting Schools

Schools were informed of the longitudinal nature of the study upon recruitment into the study in 2009. At that time, schools were informed that they would be contacted again for the first follow-up and for the high school transcript collection. All 944 schools that participated in the base-year data collection were contacted to secure participation in the first follow-up. This section describes the procedures used to contact high schools to secure their continued participation in the HSLS:09 first follow-up.

4.2.1.1 Training

Six institutional contactors (ICs) were hired to work on the first follow-up recruitment task. The ICs initially received 3 days of training in January 2011, which focused on recruiting schools for participation in the HSLS:09 first follow-up, identifying a school coordinator (SC), and coordinating logistics for the in-school session. A fourth training day was held in the fall of 2011, which focused on the student enrollment status update collection and testing of the CD that launches the mathematics assessment and student survey (Sojourn CD). The training agenda is shown in figure 4.

Figure 4. Institutional contactor training agenda

Day 1	Day 2	Day 3	Day 4
• Introductions • Study Overview • Recruiter Responsibilities • District Notifications • Recruiting Schools • Working in the Institutional Contacting System (ICS)	• Data Collection Logistics • Principal Contacts • Working with the School Coordinator • Responding to Frequently Asked Questions (FAQs) • ICS Practice	• Continued Role Play and ICS Practice • Certification	• Enrollment Status Update • Introduction to Sojourn • Sojourn Testing • Responding to FAQs • Recording Information in the ICS

4.2.1.2 Notification of School Districts

Recruitment comprised a two-step process of notifying school districts and securing continued cooperation from schools. State departments of education were informed about the study prior to base-year recruitment and were not notified again in advance of the first follow-up. School district recruitment occurred prior to contacting schools for the base year. School districts not requiring formal applications were mailed a letter in January 2011. The purpose of this letter was to notify school districts that base-year participating schools would be contacted to gain permission to conduct first follow-up data collection activities. No telephone follow-up or other recruitment activities were conducted with these school districts. Approval was obtained from each of the 35 districts requiring research application renewals prior to contacting schools within the districts.

4.2.1.3 Securing Renewal of School Cooperation

Prior to mailing letters to participating base-year schools, ICs conducted searches on the Web to confirm that the current school principal and base-year SC were still at the schools. Approximately 1 week after the district notification letter was mailed, or upon approval of a research application, if applicable, recruitment materials were sent to principals of base-year participating schools. RTI contacted school principals by telephone to discuss study details, to answer questions, and to secure cooperation for the school’s continued participation in HSLS:09. In-person visits were attempted to secure cooperation from eight schools, six of which participated in the first follow-up.

RTI asked each school to designate an SC for the study and checked to see if the base-year SC was available to reprise his or her role. The SC worked with RTI to handle logistical arrangements, which included scheduling a date for the in-school session, specifying the type of parental permission required for the in-school student sessions (explicit [active] consent or implicit [passive] consent), and obtaining the names of the staff who should receive the school administrator and school counselor questionnaires. The SC also informed the recruiters who would handle the responsibility of providing enrollment status update information for the

students sampled in the base year. An information technology (IT) coordinator was also identified to assist with testing of the Sojourn CD, which was enhanced from the base year.

As a token of appreciation for their continued participation, schools were given the option of receiving various science-based magazine subscriptions. The options provided to schools were a 1-year subscription to *Science News*, or a 2-year subscription to both *Discover* and *Popular Science* magazines. Schools already subscribing to those magazines were given a Staples gift card instead. Magazine subscriptions commenced at the start of the 2012–13 school year. Each of the 904 participating schools was offered the magazine subscriptions, of which 888 opted for the magazine subscriptions and 16 opted for the Staples gift card.

4.2.1.4 Enrollment Status Update

Schools were asked in the fall of 2011 to provide an enrollment status update for sampled students. This activity helped in planning for the first follow-up data collection because out-of-school data collection activities (e.g., parent interviews, student surveys and assessments not conducted within the school) could commence upon the start of data collection for students no longer enrolled at the base-year school. Schools were asked to provide information about each sampled student, including whether the student was still enrolled at the school and the student's current contact information. For students who were still enrolled at the base-year school, schools were asked to provide the student's current grade level in addition to family contact information. If a student was no longer enrolled, the school was asked to provide information about why the student left the school, in addition to the student's last date of attendance.

Reasons for leaving the base-year school included transferring to a different school, dropping out of school, graduating early, becoming homeschooled, leaving the country, or becoming institutionalized. If the student transferred to another school, schools were asked to provide the name, city, and state of the transfer school. Dropping out of school was defined as an absence of 4 weeks or more not resulting from illness or injury. Updated contact information was also requested for each student as part of the enrollment status update request. ICs prompted for outstanding enrollment status updates during conversations to coordinate logistics for the in-school session. If the school did not provide the information in advance of the in-school session, the SA collected enrollment status information during initial conversations with each school.

Based on the information provided from base-year schools, additional schools were contacted to conduct small group sessions if four or more students transferred to the same school that was not in the base-year HSLS:09 sample. Fourteen such schools were identified, eight of which participated in the first follow-up. Results of the student enrollment update can be found in section 4.4.2.

4.2.2 Tracing the Student Sample

HSLS:09 sample members had not been contacted since the conclusion of the base year in 2010. Three activities were performed to trace the sample in advance of the first follow-up

data collection: (1) parents were sent address update requests to ensure that contacting information for each sample member was current; (2) the sample was sent for batch tracing; and (3) schools were asked to provide updated family contact information as part of the enrollment status update.

Parents were sent a letter in September 2011, which informed them that the first follow-up data collection would begin in January 2012. The letter also requested that they log on to the HSLS:09 website to update contact information for the student and at least one parent. A paper version of the address update form and a business reply envelope were also included with the letter so parents could respond by mail in lieu of the website.

In October 2011, all cases were sent to Accurint and FirstData for batch tracing prior to the first follow-up data collection. Batch tracing service providers use proprietary databases to verify a sample member's locating information. If current contact information did not match, the tracing service delivered new contact information. All new contact information was uploaded to the first follow-up locator database prior to data collection.

4.3 First Follow-up Data Collection Methodology

This section documents the data collection methods and training protocols employed for the HSLS:09 first follow-up, including in-school and out-of-school student data collection, parent data collection, and staff data collection. Data collectors were trained to perform each component of the data collection. Session administrators (SA) performed the in-school and field data collection components and were trained over a period of 5 days to perform this work. Telephone interviewers (TIs) were trained to contact students and parents by telephone.

4.3.1 Student Data Collection Procedures

Student data collection was conducted in 904 of the 939 high schools from January 31, 2012, through June 14, 2012, with field and telephone follow-up continuing through October 28, 2012. Trained SAs conducted the in-school student sessions and field follow-up. Telephone follow-up was conducted by trained TIs.

Student data collection comprised a 35-minute questionnaire and a 40-minute mathematics assessment. Students who were still enrolled at the base-year school at the time of the first follow-up were asked to participate in their school. Students who missed the in-school session and those who were no longer enrolled at the base-year school were contacted to participate via Web, computer-assisted telephone interview (CATI), or computer-assisted personal interview (CAPI).

4.3.2 Session Administrator Training

In January 2012, approximately 140 field supervisors (FSs) and session administrators (SAs) were hired and trained. FSs and SAs were asked to review the field manual and to complete a home study exercise before the in-person training. As shown in figure 5, the training covered both the in-school and field components of data collection. In addition to preparing the SAs to conduct the student sessions, the training also covered preparing for the student sessions. This included managing the parental consent process, verifying student enrollment status, and overseeing logistical preparations in advance of the student session. The field data collection training focused on locating and contacting students and parents in the field, gaining parent permission when needed, conducting interviewer-administered questionnaires, and tracing sample members.

Figure 5. Session administrator training agenda

Introduction	Day 1	Day 2	Day 3	Day 4	Day 5
• Welcome and introductions • Overview of HSLS:09 • Prior longitudinal studies • Confidentiality and data security • Sojourn demo • Training objectives	• Classroom introductions • Purpose, background, and school administrator responsibilities • Recruiting schools • Student enrollment verification • Parental permission • Laptop overview and equipment security • Managing in-school assignments • Session logistics • Working with the school coordinator • Case Management System (CMS)	• Q&A • Sojourn practice • Questionnaire capability, eligibility, exclusions • Student tracking form • Student script • Student questionnaire and assessment prototype • Parent and staff questionnaires • Contacting parents for session reminders or refusal conversion • Reporting to the field supervisor (FS) • Test day summary form (TDSF)	• Questions from previous days • Dealing with disruptive students and problems at school • Honoraria and incentives • E-mail and transmission • Introduction to out-of-school activities • Field case incentives • Obtaining parent permission for students participating out of school • Gaining cooperation and refusal conversion • Non-interview field cases • Tracing and locating sample members	• Questions from previous days • Field (CMS), and Case Assignment Card • Event codes, review and exercise • Round robin, paired mocks • Reporting to FS on out-of-school activities • General administrative procedures	• Questions from previous days • In-school review • Out-of-school review • Q&A/ certifications

Before commencing work on the study, each SA was required to pass a series of certification assessments to demonstrate mastery in his or her job duties

SAs recruited, hired, and trained session administrator assistants (SAAs) to help during the in-school sessions, as needed. The SAA was responsible for helping set up the school computers and monitoring the student sessions. SAAs were most often used to assist with the SA’s first assigned school, when five or more laptop computers had to be carried into the school and required monitoring, and when schools split the students into multiple computer labs for concurrent sessions.

4.3.3 In-school Student Data Collection

The SA took over responsibility for the school from the ICs approximately 4 weeks prior to the scheduled session. For each assigned school, the SA was given a School Logistic Form (SLF), which was an electronic form that contained general information about the school and the logistics arranged for the session. The SLF contained information such as the name and phone number for the SC, date and time for the first follow-up testing session, special instructions regarding the school's computer lab, and logistic information from the base-year data collection. Additionally, the SAs were provided with a Student Tracking Form (STF) for each of their assigned schools. The STF was a hardcopy form which contained the roster of sampled students currently enrolled at the school. The SAs used this form to prepare for the session; confirm student enrollment; and track parental permission status, student questionnaire capability,[26] and completion codes.

During the weeks leading to the in-school session, the SA periodically contacted the SC to confirm session logistics, review student enrollment and eligibility, and ensure that consent forms were distributed and collected, as applicable. During these contacts, the SAs asked the SCs to provide updates to the students' enrollment statuses and to determine the questionnaire capability of students on the STF.

Most contacts with the SC before the in-school session were made by telephone with the exception of an in-person visit that was typically scheduled 2 weeks before the in-school session date. During this visit, the SA tested the Sojourn CD and thumb drives on school computers and assessed any additional logistical needs for the session. If the SA needed assistance at a particular school, SAs recruited, hired, and trained an assistant to help during the in-school session.

On the session date, SAs arrived at least an hour prior to the session start time to prepare the room. School computers were booted up with Sojourn, and project laptop PCs were set up as backup. If a school did not use Sojourn at all, the SA brought additional laptop PCs to schools. After reviewing parent permission status for each student, the SA took attendance, distributed user name and password information, and read the informed consent script. During the session, the SA monitored the classroom and addressed questions and technical issues, as needed.

4.3.3.1 Parent Permission

Additionally, the SAs determined whether the SC had received any parental refusals. If so, the SA began refusal conversion efforts if the school permitted. In explicit permission schools, the SC also informed the SA which parents had not yet returned the permission form. If the school permitted, the SA made prompting calls to ask these parents to return the form.

[26] Most students were questionnaire capable. Students who were classified as questionnaire incapable were those who had a physical or cognitive disability that prohibited their participation, or students who had not had at least 3 years of English instruction and were not proficient in English.

Table 17 shows the parental permission requirements for schools in the base year and first follow-up. Each school had a choice to use either explicit (written) or implicit permission forms to obtain parental permission to participate. Most schools opted to use the same type of parental permission as was used by the school in the base year. Fifty-four schools that required written permission in the base year switched to implied permission for the first follow-up. Seven schools switched from implied permission in the base year to written permission in the first follow-up. Table 18 shows the in-school response rates by permission requirements for students in both the base-year and first follow-up data collections.

Table 17. Permission requirements of base-year and first follow-up schools: 2012

	Base year		First follow-up	
Permission type	Number of schools	Unweighted percentage	Number of schools[1]	Unweighted percentage
Total	944	100.0	904	100.0
Explicit permission	190	20.1	135	14.9
Implicit permission	754	79.9	769	85.1

[1] The HSLS:09 school sample included all schools that participated in the base-year data collection. However, four schools were closed and one school did not have any base-year students still enrolled. Of the 939 sampled schools, 904 conducted student sessions in school.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Table 18. In-school student response rates, by permission requirements for base-year and first follow-up data collections: 2012

	Base year			First follow-up		
Permission type	Eligible students[1]	In-school respondents	Unweighted percentage	Eligible students[1]	In-school respondents	Unweighted percentage
Total	25,206	21,001	83.3	17,914	16,583	92.6
Explicit permission	5,090	3,327	65.4	1,298	493	38.0
Implicit permission	20,116	17,674	87.9	16,616	16,090	96.9

[1] Only includes students identified by their schools as currently enrolled as of the test date.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

4.3.3.2 Student Eligibility and Special Accommodations

Study procedures were designed to include as many students sampled in the base year as could be validly assessed or surveyed. Students were eligible to participate even if they did not participate in the base year or if they were not capable of participating in the base year. As in the base year, the SAs worked with the SCs to determine the eligibility and capability status of each sampled student. Because eligibility had been established in the base year, students were generally considered eligible unless they had not been in ninth grade during the base year (in

other words, sampled in error based on student list data provided in the base year, but only discovered to be sampling errors in the first follow-up).

In addition to determining each student's eligibility status, the physical and cognitive ability of each student to participate in HSLS:09 was also assessed. Students were included in HSLS:09 if they were physically or cognitively capable of participating in other standardized tests. Students who had been classified as English Language Learners (Limited English Proficient) during the base year would have had at least 2 additional years of English language instruction at the start of the first follow-up data collection. As a result, all but four of the 207 students who had been identified in the base year as questionnaire-incapable because of a language barrier were now questionnaire-capable. An additional 247 students in the first follow-up were identified as questionnaire-incapable because of a physical or cognitive disability for a total of 251 questionnaire-incapable students. The breakout of student eligibility and questionnaire capability is shown in table 19.

Table 19. Student eligibility and questionnaire incapability rates: 2012

	Base year		First follow-up	
	Number	Percent	Number	Percent
Total	26,305	100.0	25,206	100.0
Eligible	25,206	95.8	25,184	99.9
Questionnaire capable	24,658	97.8	24,933	99.0
Questionnaire incapable[1]	548	2.2	251	1.0
Questionnaire incapable—physical disability	38	6.9	38	15.1
Questionnaire incapable—cognitive disability	303	55.3	209	83.3
Questionnaire incapable—English Language Learner	207	37.8	4	1.6
Ineligible[2]	1,099	4.2	22	0.1

[1] Percent of total eligible students.

[2] Students were ineligible for HSLS:09 if they were not in ninth grade during the base-year data collection, they were not enrolled at the sampled high school during the base year, or if they were foreign exchange students. Deceased students as of the first follow-up are included in this number.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Special accommodations were provided to students who were questionnaire-capable, but would not have been otherwise able participate in the first follow-up. For example, students with learning disabilities or visual impairment could have someone read the questionnaire aloud to them. Students who were assisted by a reader were only eligible to take the questionnaire; the mathematics assessment could not be read to a student because of the nature of the questions (e.g., mathematical symbols, interpretation of graphs). Sign language interpreters, if provided by the school, were permitted to sign the testing instructions to students with hearing impairments. Students were given extra time on the questionnaire, the assessment, or both, if they had an

Individualized Education Plan that made such a stipulation. As shown in table 20, a total of 176 in-school student respondents (0.9 percent of the 20,594 student participants) required accommodations to participate.

Table 20. Accommodations for participating students: 2012

	Base year		First follow-up	
	Number	Percentage	Number	Percentage
Total	456	100.0	176	100.0
Extra time on questionnaire or test	381	83.6	138	78.4
Other accommodations	75	16.5	38	21.6

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Students were allotted 35 minutes to complete the questionnaire. After 35 minutes (or upon completion of the survey if completed in less than 35 minutes), students automatically transitioned to the 40-minute mathematics assessment. If the student completed all of the assessment items before the 40 minutes elapsed, but had not yet completed all the items on the questionnaire, the student cycled back to the questionnaire to answer any remaining questions. This feature was used by 2,634 students who responded to at least one questionnaire item after completing the assessment. Of those who cycled back to work on the student questionnaire, all students ultimately completed the questionnaire.

Makeup sessions were scheduled for students who missed the initial in-school session. Students who missed all in-school sessions were contacted to participate out of school. When allowed by the school, a $10 incentive was paid to students after completing the questionnaire and assessment in school.

The SC role involved coordinating logistics, disseminating and collecting permission forms, preparing enrollment status information or coordinating with the registrar to obtain this information, and facilitating the in-school session and subsequent makeup session on the scheduled date(s). On average, SCs spent 8 hours helping to prepare for the study. As a token of appreciation, SCs received a base honorarium of $100 with the opportunity to earn up to $50 more based on student response at the school.

4.3.4 Out-of-school Student Data Collection

Student sample members who had left the base-year school or who did not participate in the in-school session were contacted out of school. Student nonparticipation was a result of the student being absent from school or engaged in a conflicting activity on the in-school test day(s), or unwillingness of some schools to offer a makeup session. Students who had dropped out of school, transferred, graduated early, or became home-schooled became part of the out-of-school data collection at the start of the data collection period. Students who missed the in-school

session were added to the out-of-school data collection on a flow basis as each school's in-school activities concluded. Unlike the base year, the algebraic reasoning assessment was made available to out-of-school student respondents in the first follow-up.

4.3.4.1 Telephone Interviewer Training

Telephone interviewer staff for the first follow-up data collection included quality control supervisors (QCSs), help desk agents, TIs, and intensive-tracing staff. Before working on HSLS:09, all staff completed a comprehensive training regimen. As detailed in figure 6, the trainings covered helping sample members troubleshoot and resolve problems, completing the questionnaire over the Web, locating and interviewing parents and students in CATI, gaining parent permission when needed, and prompting students to complete the math assessment. All trainees were required to pass certification assessments associated with the questionnaires, the CATI Case Management System (CATI-CMS), and frequently asked questions.

Figure 6. Web/CATI training agendas

Supervisor training	Help desk training	Telephone interviewer training
Welcome and Introduction	Welcome and Introduction	Welcome and Introduction
Overview of Study	Overview of Study	Overview of Study
Confidentiality	Confidentiality	Confidentiality
Your Role as a Supervisor	Your Role as a Help Desk Agent	Your Role as a Telephone Interviewer
Monitoring	Frequently Asked Questions	Frequently Asked Questions
Incident Reports	Parent QxQ Review	Front-end Overview
Front-end Overview	Parent Paired Mock	Front-end Practice
Case Management	Front-end Overview	Help Desk Demonstration
Case Review	Front-end Practice	Coding Exercises
	Student QxQ Review	Parent Round Robin
	Student Paired Mock	Student Round Robin
	Coding Exercises	Front-end Practice
	Help Desk Demonstration	Frequently Asked Questions Practice
	FAQ Practice	Parent Paired Mock
	Additional Coding Practice	Additional Coding Practice
	Parent Paired Mock	Student Paired Mock
	Student Paired Mock	Monitoring and Supervision
	Training Evaluation	Training Evaluation

4.3.4.2 Contacting and Interviewing

The first phase of the out-of-school data collection—a 3-week web-based early data collection period—began after the initial contact mailing was sent to sample members in February 2012. The initial contact mailing consisted of a study brochure and a letter which included the HSLS:09 website address, a unique study ID and password, and a request that the student complete a self-administered questionnaire on the Web. Mail and e-mail reminders were

sent to nonresponding sample members every 3 weeks. At the conclusion of the early data collection period, telephone interviewers (TIs) began placing outbound CATI calls to nonresponding students to request that they complete their parent questionnaire over the telephone. Sample members could still complete a self-administered Web questionnaire until data collection concluded on October 28, 2012.

Parent permission was required before students under the age of 18 could participate in HSLS:09. Four methods were used to obtain parent permission. First, when permission was required, all parent contacting materials included a sealed envelope with the student letter. The parents were instructed to give the enclosed letter to the students and that by doing so they granted their permission to the student. Second, parent contacting materials instructed parents to log into the study website to provide their permission online, as needed. Third, parents were prompted to give their permission after logging into the parent questionnaire. Fourth, TIs prompted parents prior to launching the parent questionnaire. When permission was obtained in this manner, the interviewer attempted to conduct the student questionnaire during the same call. After parent permission was granted (or if the student was 18 or older), all phone calls, letters, and e-mails regarding the student survey were directed to the student, until the student questionnaire was completed.

Cases were assigned to field staff (i.e., SAs) when the contacts had been exhausted in CATI and the case was close to an SA travel area. Student and parent cases were always worked in the same out-of-school mode. For example, if a student was absent from the in-school session, the student case would go to CATI data collection if the parent case was currently being worked in CATI. Similarly, the student case would be sent to the field if the parent case was being worked in the field.

For student cases worked by field staff, the SA also provided the study information first to the parents to obtain permission to conduct the student interview. Information about contacting parents is described in section 4.4.5. Field interviewers attempted to contact sample members to conduct interviews by telephone before any in-person attempts were made. In-person contacts were attempted after telephone contact was unsuccessful.

Students who participated by Web or in person were able to move directly from the questionnaire to the assessment in the same session, which matches the experience of students who participated in school. However, students who participated by phone could complete the questionnaire over the phone but had to log into the study website to complete the mathematics assessment. Students who were still enrolled at their base-year school were offered a $15 incentive to complete the student questionnaire during the out-of-school data collection. Students who were no longer enrolled at their base-year school were offered $40 to complete the student questionnaire.

If asked to complete the mathematics assessment outside of the school setting, students were offered an additional $10. Some 3,271 assessments were completed out of school, and

15,236 were completed in school. (Chapter 2 addresses the comparability of in-school and out-of-school assessment results.)

4.3.5 Parent Data Collection Procedures

One parent or guardian of the subsample of students was asked to complete a parent questionnaire. The parent subsample comprised 11,961 parents, including 9 parents who were subsequently deemed ineligible because the associated student was ineligible or deceased. It was preferred that the parent most knowledgeable about the student's education be the person to respond to the parent questionnaire. Parents had the option of participating via Web, CATI, or CAPI. This section details the procedures for the parent data collection.

4.3.5.1 Web/CATI Procedures

Parents followed the same data collection procedures described below in section 4 with a 3-week web-based early data collection period. At the conclusion of the early data collection period, TIs began placing outbound CATI calls to nonresponding parents to request that they complete their parent questionnaire over the telephone. Parents could still complete a self-administered Web questionnaire until data collection concluded on October 28, 2012. Parents who were not able to complete a questionnaire in English were given the option to complete the interview in Spanish.

4.3.5.2 Field Procedures

Based on outcomes of the response propensity model described below in section 4, a subset of parent cases were assigned for field data collection. The cases identified as having the lowest propensity to respond were sent to the field (field intervention) to be attempted by phone or in-person by field interviewers. Field interviewers first attempted to complete their assigned cases by telephone to conserve costs. In-person visits were made after telephone attempts had been exhausted. As previously mentioned, student cases that were part of the out-of-school data collection were worked in the same mode that the parent case was being worked. For each parent case, SAs received all contact information available for the case as well as any available tracing and CATI history information to assist the SA in locating the sample member. SAs located the sample members, gained cooperation to participate in the study, and conducted the interview with the sample member.

Additional cases were assigned for field interviewing throughout the data collection period. Cases were assigned for field interviewing based on proximity to the field interviewers or travel assignments, and the number and success of call attempts in CATI. In total, 3,386 parent cases were assigned to field staff during the data collection period.

4.3.5.3 Parent Incentive and Paper-and-Pencil Procedures

Beginning in the first week in July, a $20 incentive was offered to a subset of parents who met the criteria of being a "most challenging" case. The cases designated as "most

challenging" had high call counts in CATI or in the field, had ever refused to participate in the first follow-up,[27] or had an address but no good phone number to contact. A total of 5,125 parents were offered the $20 incentive by the end of data collection.

Of the 5,125 parents that were offered the $20 incentive by the end of data collection, 3,269 parents were sent a paper-and-pencil (PAPI) questionnaire during the first week of August, 2012. The remaining 1,856 completed the parent interview or their case was otherwise finalized after being offered $20 but prior to mailing PAPI questionnaires. The PAPI questionnaire was an abbreviated version of the parent questionnaire containing nine critical items.

As part of the PAPI administration, a $5 prepaid incentive accompanied this mailing along with a business reply envelope (BRE) to facilitate its return. Therefore, the letter indicated that the parent would still receive the $20 for completing the full questionnaire. However, parents who completed the PAPI questionnaire but not the full questionnaire were not eligible for the $20 incentive and therefore received just the $5 prepay.

About 3 weeks prior to the end of the data collection period, a second PAPI questionnaire and BRE was sent to nonresponding parents. The second PAPI questionnaire did not include the $5 prepay incentive. Upon return to RTI, completed PAPI forms were processed and the data were keyed into a data entry program. Table 21 below describes the cases offered (1) neither the $20 incentive nor the PAPI questionnaire, (2) only the $20 incentive, or (3) both $20 incentive and PAPI questionnaire.

Table 21. Parent questionnaire responses, by incentive and PAPI offers: 2012

		Completed questionnaires			
		Full		PAPI	
	All parent cases	Number	Percent	Number	Percent
Total	11,952	8,176	68.4	475	4.0
No incentive or PAPI	6,827	5,950	87.2	†	†
Offered only $20	1,856	874	47.1	†	†
Offered PAPI with $5 prepay and $20 offer	3,269	1,352	41.4	475	14.5

† Not applicable.

NOTE: All percentages are based on the number of students within the row under consideration. PAPI = paper-and-pencil interview.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

[27] A sample member was designated as "ever refused" if he or she had ever refused to participate in the HSLS:09 first follow-up. Base-year participation status is not taken into account when determining "ever refusal" status. All refusal cases are contacted at least one additional time to attempt refusal conversion.

4.3.5.4 Responsive Design–based Approach to Targeting Nonrespondents

The HSLS:09 first follow-up used a responsive design methodology to strategically target cases that could potentially contribute to nonresponse bias if they remained nonrespondents. To address the possible nonresponse bias in the parent sample, a response propensity model was used to predict cases that were least likely to become respondents. The parent cases with low response propensities were then worked by field staff.

Targeting low propensity to respond cases in nonresponse follow-up is preferable to a nonstrategic approach that attempts to raise overall response rates, which potentially includes both "easy to complete" cases and "difficult to complete cases." A nonstrategic approach to increase response rates may not successfully reduce nonresponse bias, even if higher response rates are achieved (e.g., Curtin, Presser, and Singer 2000; Keeter et al. 2000). Specifically, a nonresponse follow-up protocol may increase the nonresponse bias when the protocol brings in sample members who resemble those most likely to respond (Schouten, Cobben, and Bethlehem 2009). Decreasing bias during the nonresponse follow-up depends on the approach selected to increase the response rate (Peytchev, Baxter, and Carley-Baxter 2009).

The goal of this approach was to increase the parent response rate while simultaneously reducing the variance of the response propensities in the sample, thereby minimizing nonresponse bias. To illustrate, consider that nonresponse bias can be expressed as

$$NRBias = Cov(y_i, \rho_i)/\bar{\rho}$$

where $Cov(y_i, \rho_i)$ is the covariance between the response propensity ρ_i and some survey variable y_i, and $\bar{\rho}$ is the mean response propensity. Note that $Cov(y_i, \rho_i)$ can also be written as $\tau_{y\rho}\sigma_y\sigma_\rho$, where $\tau_{y\rho}$ is the correlation between response propensity and the survey variable, σ_ρ is the standard deviation of response propensities, and σ_y is the standard deviation of y. Thus, as the variance of the response propensity is reduced through the inclusion of more respondents with differing y values, nonresponse bias will decrease even as $\bar{\rho}$ remains constant. If successful, the variability in the response propensities will be reduced as will the association between propensities and survey variables, thereby reducing nonresponse bias.

4.3.5.5 Responsive Design Selection Methodology

Response propensities for all cases in the first follow-up main study parent subsample were calculated after 3 weeks of early Web response period and the first 3 weeks of CATI data collection. This 6-week data collection period facilitated the collection of paradata (e.g., student, parent, and school characteristics; and panel maintenance results) and allowed time for sample members to complete a parent questionnaire during the less-costly early data collection period. At the conclusion of the first 6 weeks, cases in the lowest quartile of propensity scores were sent to the field to be attempted by a field interviewer, by telephone or in person. The dependent variable on which response propensities were based was a case's response outcome during these

first 6 weeks of data collection. The independent variables included the range of paradata mentioned above.

4.3.5.6 Responsive Design Model Specification

The propensity model developed for the HSLS:09 first follow-up data collection incorporated paradata and sampling frame variables to estimate likelihood of response. The use of survey variables was investigated; however, no survey variables included all of the known values for both respondents and nonrespondents. Imputation for nonrespondents on survey variables was considered, but ultimately rejected because the level of missing values exceeded 25 percent for most cases, given that the first follow-up included base-year nonrespondents as well as base-year respondents.

Both student- and parent-level variables were included in the propensity model including base-year response status for the parent and student; outcomes (response status and mode) for the first follow-up panel maintenance activity; school-provided enrollment status of the student (dropout, transfer, homeschooled, still enrolled at the base-year school); early graduate status; whether the student was a refusal or absent during the base year; call counts in the base year and first follow-up; number of contact attempts in the base year and first follow-up; whether the sample member made an appointment for the interviewer to call at another time; whether the case logged into but did not complete the Web questionnaire; sex; race; school type; type of metropolitan area; and region.

At the time of model implementation, 3,385 parent cases had responded, leaving 8,065 pending cases that were available for consideration for the field intervention (CAPI) and were assigned to field interviewers to be completed by phone or in-person. To assign propensity scores to nonresponding cases, a logistic regression model was fit to the data and predicted probabilities were used to score cases. Variables retained through backward-selection, odds ratios, and confidence intervals for all retained variables used in the propensity score calculation are listed in table 22.

Table 22. Odds ratios and confidence intervals for retained variables for responsive design case selection: 2012

Retained variable	Odds ratio	95% confidence interval
Parent refused during first follow-up	0.126	0.083, 0.191
Logged into the Web questionnaire (first follow-up) but did not complete	0.192	0.107, 0.345
Call count during first follow-up	0.958	0.955, 0.961
Lives in a rural area	1.269	1.141, 1.41
Lives in the South	1.329	1.199, 1.473
Number of first follow-up contact attempts	1.373	1.339, 1.408
First follow-up dropout	1.485	1.104, 1.996
Lives in the Midwest	1.541	1.372, 1.73
Made a soft appointment	1.705	1.535, 1.894
Was a Web respondent to the first follow-up panel maintenance activity	1.733	1.408, 2.134
In the same school in first follow-up and base year	1.777	1.588, 1.989
Student was a respondent in the base year	1.809	1.527, 2.144
Made a hard appointment	1.863	1.679, 2.068
Parent was a respondent in the base year	2.139	1.925, 2.376
Responded to the first follow-up panel maintenance activity	2.931	2.522, 3.405

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Cases in the lowest quartile of response propensities (i.e., those cases least likely to respond) were identified as cases that would be sent to the field, because the cases were least likely to respond. However, 218 of the 2,016 cases identified to be sent to the field were withheld for reasons such as the case had been finalized in CATI, the case had been started but not completed in another mode (i.e., CATI or the Web), or the case had insufficient locating information to contact the sample member in the field. In total, 1,798 cases were sent to the field as a result of the response propensity lowest likelihood determination.

4.3.6 Quality Control Procedures

Multiple strategies were used to ensure that high-quality data were collected in each mode of data collection. The strategies implemented to ensure data quality at each mode are discussed in this section.

4.3.6.1 In-school Data Collection Quality Control Procedures

Quality control procedures implemented as part of the in-school data collection included weekly individual and group meetings, in-person observations, and verification calls to the schools after the sessions were completed. Retraining was conducted as necessary as a result of these quality control measures throughout the course of data collection.

SAs had weekly meetings with their supervisors to discuss the status of the schools in their assignment. These calls were designed to talk about plans to maximize student response rates at each school, identify potential challenges and proactively brainstorm solutions, and ensure that the SA had all of the materials needed to conduct the session. In addition, weekly group calls were conducted to share information about challenges experienced in the schools and strategies that have proven successful. The SAs also touched base with their supervisor after each session to report on how things went and to discuss strategies for the makeup sessions, if necessary.

In addition, after each session, a supervisor contacted the SC to obtain feedback on how the session went overall. SCs were asked if the session started on time, if the SA was professional, and for any comments about the overall experience. ICs also called the SC once all school-related activities were completed to thank the school for its participation and to obtain any additional feedback about the process as a whole.

In-person observations were conducted unannounced at 5 percent of participating schools in the first follow-up. A member of the project team visited the school on the day of the session and observed the session taking place. Feedback was provided to the SA to discuss areas that were successful as well as opportunities for improvement.

4.3.6.2 Web/CATI Data Collection Quality Control Procedures

Quality control procedures implemented for the Web/CATI data collection included live interviewer monitoring, a help desk, and regular Quality Circle (QC) meetings.

Call center supervisory staff conducted live monitoring which allowed supervisors to oversee TIs in real time. This ensured that interviewers were following scripts, coding responses accurately, and conducting interviews in a professional manner. Additionally, RTI project staff monitored recorded interviews for quality assurance purposes. Approximately 10 percent of the cumulative interviewing hours were monitored.

HSLS:09 staffed a help desk to assist parents and students accessing the Web questionnaire, answer questions about the study, and give sample members log-in information. Sample members could reach the help desk via a toll-free telephone number or by e-mail. Both methods of contacting the help desk were provided in all materials sent to the sample members.

Sample members primarily contacted the help desk to request a new password, their study ID, or both. Help desk agents were instructed to offer to conduct the interview over the telephone if they were unable to resolve the caller’s issue within 5 minutes. Calls to the help desk were entered into a help desk application. The help desk application allowed help desk agents to

- document each call to the help desk;
- verify a caller’s identity;
- provide study IDs and passwords; and

- follow up with calls to the help desk that had not been resolved immediately.

Regular QC meetings were held to ensure that a line of communication remained open between the TIs and project staff. Meetings were held weekly during the first 2 months of data collection, and then biweekly thereafter. Notes from each meeting were posted on the call center network so interviewers and supervisors could review the notes when they were not at the meeting. Topics discussed during HSLS:09 first follow-up QC meetings included issues with the questionnaires, gaining cooperation, avoiding and converting refusals, leaving detailed case comments in CATI, help desk operations, and getting feedback about data collection issues.

4.3.6.3 Field Data Collection Quality Control Procedures

To ensure quality control, FSs contacted 10 percent of all cases completed in the field to ensure that the case was conducted appropriately and collect feedback from the sample member. FSs met with field interviewers each week to discuss progress on cases, plans to work remaining cases, and feedback from verification contacts.

4.3.7 Procedures for Staff Surveys

In addition to the student and parent questionnaires, one administrator and one counselor at each base-year school was asked to complete a 30-minute questionnaire. In addition, for students who had transferred to a different school since the base year, one school administrator from each transfer school was asked to complete an abbreviated version of the school administrator questionnaire (there was no counselor questionnaire for the transfer school). Unlike the base year, no teachers were surveyed in the first follow-up. Each staff questionnaire was available to be completed via the Web or via CATI. Toward the end of data collection, an abbreviated school administrator questionnaire was available by PAPI for transfer schools and for late responding base-year schools. The following section summarizes the training and data collection procedures for staff surveys.

4.3.7.1 Staff Data Collection Training

In addition to recruiting schools, ICs also collected data for staff data collection. The ICs prompted administrators and counselors to complete their questionnaires, served as help desk agents, and conducted staff interviews by telephone. The ICs also conducted verification calls upon completion of the school's participation to ensure that the student session at the school was successfully implemented. The ICs received one day of staff data collection training, for which an agenda can be found in figure 7.

Figure 7. Staff data collection training agenda: 2012

- Overview of IC Responsibilities/Tasks—Staff Data Collection
- Prompting/Contacting Materials
- Phone Interviews
- Help Desk
- Verification Interviews/Course Catalog Collection
- Administrator and Counselor Surveys
- Interviewing Techniques
- Review of questionnaires
- ICS/Help Desk Demo
- Verification Calls/Course Catalog Collection
- Questions & Answers

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

4.3.7.2 School Staff Collection Procedures

Staff data collection took place from February through October 2012. One administrator and one counselor from each base-year school was asked to complete the school administrator and school counselor questionnaire, respectively. During the recruitment phase, the SC was asked to identify and provide contacting information for the appropriate staff members to complete the school staff questionnaires, regardless of whether the school was participating in the in-school student component of the study. A letter and study brochure were mailed to the staff members tasked with completing the staff questionnaires. ICs prompted nonrespondents throughout the data collection period.

First follow-up administrator questionnaire for base-year schools. School administrators were asked to provide input about the administration and policies at their schools. The letter sent to the administrator provided instructions on how to access the web-based questionnaire and how to complete the questionnaire by telephone with one of RTI's ICs. The web-based questionnaire could only be accessed using the staff member's unique study ID and password. Periodic letters and e-mail reminders were used to prompt administrators to complete the questionnaire in addition to telephone prompts by the ICs. The self-administered school administrator questionnaire took approximately 29 minutes to complete.

Because the majority of the administrator questions were general questions about the school's policies and procedures, the administrator could designate a knowledgeable staff member to complete the first three sections of the questionnaire. The final section of the questionnaire was to be completed only by the principal of the school. The final section took less than 8 minutes to complete, thereby reducing the burden on the principal by allowing another staff member to complete the first three-quarters of the questionnaire.

An abbreviated administrator questionnaire, which could be completed via the Web or over the telephone, was made available to the schools during the last few weeks of data collection. This questionnaire asked critical items from the full administrator questionnaire,

including items such as school type, enrollment, course scheduling, and number of teachers at the school. The abbreviated administrator questionnaire was designed to be completed by either the school principal or another knowledgeable staff member and took approximately 9 minutes to complete. A hardcopy version of the abbreviated administrator questionnaire was also mailed to nonresponding school administrators.

First follow-up administrator questionnaire for new schools (transferred to by the dispersing cohort). The transfer school administrator data collection took place from June through October 2012. Collecting data from transfer school administrators provides data that have not been collected in prior longitudinal secondary school studies. Transfer schools were identified on an ongoing basis from the enrollment status information provided by base-year schools, as well as by students in their first follow-up questionnaires. Transfer schools were only contacted if at least one student who transferred to the school participated in the student component of the study. Contact information for each transfer school was obtained using NCES's Common Core of Data (CCD) and school and district websites.

A letter and study brochure were mailed directly to each transfer school administrator requesting the administrator to complete the abbreviated school administrator questionnaire via the Web which took approximately 8 minutes to complete. The abbreviated school administrator questionnaire given to transfer school administrators was the same abbreviated form offered to base-year school administrators as described above. The letter also included the study telephone number for the administrator to use in case of questions or to complete the questionnaire by phone. In addition to telephone prompts made by the ICs, letters and e-mail reminders were periodically sent to prompt administrators to complete the questionnaire. A hardcopy version of the questionnaire was also mailed to nonresponding school administrators a few weeks before the end of data collection.

Counselor questionnaire. School counselors were asked to provide input about student placement into classes, counselor resources available to students, graduation requirements, and college preparation programs offered at the school. The ICs collaborated with the SC to name the counselor to complete the questionnaire, and typically the lead counselor was named. A lead letter and a study brochure were sent to the person identified to complete the school counselor questionnaire. The letter provided instructions on how to access the web-based questionnaire and how to complete the questionnaire by telephone with an IC. The web-based questionnaire, which took approximately 29 minutes to complete, could only be accessed using the staff member's unique study ID and password. Periodic letters, e-mail reminders, and telephone prompts were used to prompt counselors to complete the survey.

4.4 Data Collection Results

The following sections provide results for the HSLS:09 first follow-up data collection. Starting with school retention and enrollment status, results are provided for the student

questionnaire and mathematics assessment as well as the parent, school administrator, and counselor questionnaires.

4.4.1 School Retention

School recruitment commenced with 944 schools that had participated in the base year. Table 23 displays the school sample sizes among specific sampling strata and participation yield by school type and locale. Five of the 944 schools were found to be closed or had no eligible sampled students still enrolled in the base-year school. Of the eligible 939 schools, 904 base-year schools (96 percent) agreed to continue participation in the HSLS:09 first follow-up. Schools declining to conduct student sessions in the first follow-up were nevertheless asked to complete the school administrator and school counselor questionnaires. Sampled students who attended schools that did not conduct in-school sessions in the first follow-up were contacted to participate outside of school.

Table 23. School participation rates, by selected base-year school characteristics: 2012

School characteristic[1]	Eligible schools	Participated Number	Percent
Total[2]	939	904	96.3
School Type			
Public	765	740	96.7
Total private	174	164	94.3
Catholic	98	95	96.9
Other private	76	69	90.8
Locale			
City	270	259	95.9
Suburban	334	320	95.8
Town	116	111	95.7
Rural	219	214	97.7
Region			
Northeast	148	143	96.6
Midwest	249	239	96.0
South	379	372	98.2
West	163	150	92.0

[1] School characteristics shown here were updated after the base-year data collection and thus may be different than those reported in chapter 3.

[2] The HSLS:09 first follow-up school sample included all schools that participated in the base-year data collection. However, four schools were closed and one school did not have any base-year students still enrolled.

NOTE: All percentages are based on the number of students within the row under consideration.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year.

4.4.2 Enrollment Status

Enrollment status information was collected at two time points in the data collection period. Each school participating in the first follow-up was asked to provide the enrollment status of the students sampled from the school. This was provided in the months leading up to the data collection or the SAs collected updated enrollment status information from schools in the weeks leading up to the in-school session if it had not already been received. As part of the student questionnaires, students also provided their current school situation. School-provided enrollment status was used to prepare for the data collection. It was not until the student questionnaire was completed that it was possible to determine whether there were discrepancies between the school and student reports. In about 8 percent of cases, the school either did not know what the student was doing after leaving the base-year school or did not provide enrollment status for a particular student who also did not participate in the in-school session.

Table 24 shows enrollment status as provided by both schools and students. School-provided enrollment status includes all sampled students while student-provided enrollment status comprises only those students who participated in the first follow-up. Schools reported that about three-quarters of the sampled students remained at the base-year school 2.5 years later, and 84 percent of student participants were enrolled at the base-year school. Schools reported that most students who left the base-year school transferred to another school. Schools were unable to provide information about the enrollment status of 2 percent of students who left the base-year school other than to say that they were no longer enrolled at the base-year school. Enrollment status was not provided for 6 percent of the students. Given that enrollment status is unknown for about 8 percent of the sample, it is possible that there are some "hidden" dropouts. Nonetheless, to the extent that "status unknown" students participated in the first follow-up, each was asked about their 2011–12 enrollment status (i.e., whether they were dropouts).

Table 24. Enrollment status updates, by type: 2012

	School-provided enrollment status		Student-provided enrollment status	
Enrollment status	Number	Percent	Number	Percent
Total students	25,206	100.0	20,594	100.0
In base-year school	18,852	74.8	17,332	84.2
Transfer	3,635	14.4	2,391	11.6
Dropout	556	2.2	427	2.1
Early graduate	38	0.2	210	1.0
Homeschooled	219	0.9	234	1.1
Not in base-year school/reason unknown	457	1.8	†	†
Unknown status	1,427	5.7	†	†
Ineligible	22	0.1	†	†

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

4.4.3 Student Data Collection Results

HSLS:09 first follow-up student questionnaires were completed in one of four data collection modes: in-school, Web, CATI, and CAPI. The student questionnaire was completed by 82 percent of eligible sampled students in the first follow-up. Sixty-one percent of students completed the questionnaire in school, while 20 percent completed the questionnaire outside of school, which comprised students who were no longer enrolled in the base-year school and those who missed the in-school session. During out-of-school data collection, 9 percent of student respondents completed the questionnaire via the Web, 6 percent completed the questionnaire with a field interviewer, and 5 percent completed the questionnaire by phone. First follow-up student response rates are summarized by selected characteristics (including enrollment status) in table 25. Response rates by in-school and out-of-school status are presented in table 26.

Student dropout status was determined during the HSLS:09 first follow-up data collection by data provided on the student questionnaire, the parent questionnaire, and enrollment status information provided by the school. Students were defined as current dropouts if they had been absent from school for 4 or more consecutive weeks, not due to accident, illness, or vacation. Students who dropped out of high school and subsequently obtained a General Educational Development (i.e., a GED), or other high school equivalency degree, were not defined as dropouts, but rather as early graduates.

A total of 670 students (2.7 percent) was identified as current dropouts as of the spring of 2012. Of these 670 dropouts, 66 percent completed a student questionnaire. An additional 1,395 students (5.7 percent) were identified as having previously left school for 4 or more consecutive weeks not due to accident, illness, or vacation. Approximately 82 percent of the students who dropped out and returned to school completed a student interview. Overall, 2,065 students of the 25,184 eligible students (8.2 percent) were current dropouts or had a dropout episode during their education.

Table 25. Summary of HSLS:09 first follow-up student response rates, by selected characteristics: 2012

Characteristic	Response rate: Weighted[1]	Response rate: Unweighted[2]
Total	82.0	81.8
Participated		20,594
Eligible sample		25,184
Enrollment status[3]		
In base-year school	90.3	90.3
Transfer	60.3	60.8
Dropout	57.1	57.7
Early graduate	34.6 !	55.3
Homeschooled	62.5	58.9
Not in base-year school/unknown reason	54.0	54.1
Unknown status	48.9	44.9
School type		
Public	82.0	81.9
Total private	81.7	81.2
Catholic	84.0	83.2
Other private	79.5	78.2
Locale		
City	81.4	80.9
Suburban	80.2	80.6
Town	83.6	82.3
Rural	83.5	83.7
Region		
Northeast	80.9	80.7
Midwest	82.2	82.5
South	83.1	82.8
West	80.8	79.5
Sex[4]		
Male	80.8	80.9
Female	83.4	83.1

See notes at end of table.

Table 25. Summary of HSLS:09 first follow-up student response rates, by selected characteristics: 2012—Continued

Characteristic	Response rate	
	Weighted[1]	Unweighted[2]
Race/ethnicity[4]		
American Indian or Alaska Native	74.2	75.1
Asian[5]	86.1	82.7
Black, non-Hispanic	80.8	81.0
Hispanic or Latino	81.6	80.7
White, non-Hispanic	83.1	83.2
More than one race	59.7	57.9

! Interpret data with caution. Estimate is unstable because the relative standard error represents more than 30 percent of the estimate.

[1] Weighted percentages use the student base weight.

[2] Unweighted percentages are based on 25,184 eligible student sample members.

[3] Enrollment status provided by the school.

[4] Sex and racial/ethnic designations determined by data provided by (in order of preference) the sample member, parents, or school, or, in rare instances, have been imputed.

[5] The Asian category also includes Native Hawaiians and other Pacific Islanders.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 26. Summary of HSLS:09 first follow-up in-school and out-of-school student response rates, by selected characteristics: 2012

	In base-year school		Not in base-year school	
Characteristic	Weighted[1]	Unweighted[2]	Weighted[1]	Unweighted[3]
Total	90.3	90.3	57.0	56.3
Participated		17,026		3,568
Eligible sample		18,851		6,333
School type				
Public	90.4	90.4	57.3	58.4
Total private	89.7	89.9	53.3	41.5
Catholic	90.3	90.4	46.4	37.9
Other private	89.0	89.2	56.6	44.3
Urbanicity				
City	89.9	89.5	60.5	56.8
Suburban	88.4	89.0	54.7	55.5
Town	93.3	92.2	55.7	55.0
Rural	91.3	91.7	55.8	57.5
Region				
Northeast	88.8	88.8	50.4	48.4
Midwest	91.7	91.1	54.5	56.3
South	90.8	90.8	59.1	57.7
West	89.5	89.4	60.2	58.6
Sex[4]				
Male	89.9	90.1	55.0	53.8
Female	90.8	90.6	59.6	60.0
Race/ethnicity[4]				
American Indian or Alaska Native	94.6	92.8	41.4	46.9
Asian[5]	91.7	90.1	54.6	50.8
Black, non-Hispanic	91.6	91.1	58.1	60.7
Hispanic or Latino	90.2	91.4	61.6	55.2
White, non-Hispanic	90.4	90.5	56.4	58.1
More than one race	74.5	72.4	32.3	31.8

[1] Weighted percentages use the student base weight.

[2] Unweighted percentages are based on 18,851 student sample members who were still enrolled at the base-year high school.

[3] Unweighted percentages are based on 6,333 student sample members who were no longer enrolled at the base-year high school.

[4] Sex and racial/ethnic designations determined by data provided by (in order of preference) the sample member, parents, or school, or, in rare instances, have been imputed.

[5] The Asian category also includes Native Hawaiians and other Pacific Islanders.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 27 displays the time required to complete a student questionnaire. Overall, the average time to complete a student questionnaire was 33 minutes. The time to complete varied by data collection mode. In-school questionnaires averaged approximately 29 minutes, whereas CATI questionnaires took, on average, almost 49 minutes to complete.

Table 27. Average time in minutes to complete student questionnaire, by data collection mode: 2012

Mode	Average time in minutes
Total	32.5
In-school	29.3
Out of school[1]	
Web	38.3
Computer-assisted telephone interview	48.7
Computer-assisted personal interview	42.2

[1] Students who left the base-year school received different questions than students who were still enrolled at the base-year school.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 28 shows the assessment coverage of students who completed a questionnaire. Of the students who completed a questionnaire, 90 percent also completed an assessment. The assessment was completed by 98 percent of students who participated in school. The student assessment was completed by 64 percent of students who completed the questionnaire outside of school. Students who completed the questionnaire on the phone were asked to complete the assessment via the Web because of the inability to read tables and graphs over the phone. By including the out-of-school assessment in the HSLS:09 first follow-up design, student coverage was increased from 74 percent to 90 percent. Table 29 displays weighted and unweighted student assessment coverage rates by selected student characteristics.

Table 28. Student assessment coverage, by questionnaire response mode: 2012

Mode	Student questionnaire responses	Student assessments responses	Unweighted assessment coverage
Total	20,594	18,507	89.9
In-school	15,473	15,236	98.5
Out of school	5,121	3,271	63.9
Web	2,267	1,892	83.5
Computer-assisted telephone interview	1,312	508	38.7
Computer-assisted personal interview	1,542	871	56.5

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 29. Student assessment coverage, by student characteristics: 2012

Characteristics	Student questionnaire responses	Student assessments responses	Weighted assessment coverage[1]	Unweighted assessment coverage
Total	20,594	18,507	89.0	89.9
School type				
Public	17,164	15,312	88.9	89.2
All private	3,430	3,195	91.0	93.1
Catholic	2,058	1,945	92.6	94.5
Other private	1,372	1,250	89.3	91.1
Enrollment status[2]				
In base-year school	17,334	16,426	94.3	94.8
Transfer	2,393	1,553	62.8	64.9
Dropout	442	263	53.0	59.5
Early graduate	189	120	61.8	63.5
Homeschooled	236	145	59.3	61.4
Sex[3]				
Male	10,384	9,266	87.9	89.2
Female	10,210	9,241	90.1	90.5
Race/ethnicity[3]				
American Indian or Alaska Native	187	158	83.9	84.5
Asian[4]	2,129	1,988	93.6	93.4
Black, non-Hispanic	2,518	2,163	86.7	85.9
Hispanic or Latino	3,193	2,803	85.6	87.8
White, non-Hispanic	12,217	11,091	90.6	90.8
More than one race	350	304	88.5	86.9

[1] Weighted percentages use the student base weight.

[2] Enrollment status provided by the student.

[3] Sex and racial/ethnic designations determined by data provided by (in order of preference) the sample member, parents, or school, or, in rare instances, have been imputed.

[4] The Asian category also includes Native Hawaiians and other Pacific Islanders.

NOTE: All percentages are based on the number of students within the row under consideration.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

4.4.4 Parent Data Collection Results

Among the subsample of parents contacted to participate in the HSLS:09 first follow-up, about 72 percent completed a questionnaire. Table 30 shows the average time to complete a parent questionnaire by data collection mode. The average time to complete a parent questionnaire across all data collection modes was 37 minutes. Time to complete the parent questionnaire varied by mode with Web respondents averaging 34 minutes, CAPI respondents averaging 37 minutes, and CATI respondents averaging 40 minutes.

Table 30. Average time in minutes to complete parent questionnaire, by data collection mode: 2012

Mode	Average time in minutes
All parent respondents[1]	36.6
Web	33.8
Computer-assisted telephone interview	40.0
Computer-assisted personal interview	37.2

[1] Average time does not include parents who responded via paper-and-pencil interview questionnaire.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 31 displays parent participation by sample release, response propensity, incentive, questionnaire language, and data collection mode. About 53 percent of cases offered an incentive, based on many unsuccessful prior contact attempts, a refusal, or not having a telephone number for contacting, participated. Four percent of participating parents completed the questionnaire in Spanish. This table also shows the data collection period when the parent participated, and demonstrates that 65 percent of parent respondents participated at least 8 weeks or later into the data collection period during nonresponse follow-up activities.

Table 31. Parent participation by data collection phase and methodological variables: 2012

Method	Eligible parents in subsample: Total eligible parents in subsample	Eligible parents in subsample: Number of responses	Eligible parents in subsample: Percent response	Early Web: Number	Early Web: Percent of responses	Early CATI: Number	Early CATI: Percent of responses	Nonresponse follow-up: Number	Nonresponse follow-up: Percent of responses
Total	11,952	8,651	72.4	1,198	13.8	1,753	20.3	5,700	65.9
Sample release									
First release	11,441	8,310	72.6	1,123	13.5	1,692	20.4	5,495	66.1
Second release	511	341	66.7	75	22.0	61	17.9	205	60.1
Incentive offered									
Offer of $20 incentive	5,125	2,701	52.7	†	†	†	†	2,701	100.0
No incentive	6,827	5,950	87.2	1,198	20.1	1,753	29.5	2,999	50.4
Questionnaire language[1]									
English	†	8,312	†	1,194	14.4	1,742	21.0	5,376	64.7
Spanish	†	339	†	4	1.2	11	3.2	324	95.6
Mode[2]									
Web	†	3,852	†	1,015	26.3	862	22.4	1,975	51.3
CATI	†	2,955	†	183	6.2	891	30.2	1,881	63.7
CAPI	†	1,369	†	†	†	†	†	1,369	100.0
PAPI	†	475	†	†	†	†	†	475	100.0

† Not applicable

[1] Questionnaire language and mode were determined after parent questionnaires were completed.

NOTES: Results are provided for parent subsample, regardless of student participation. CAPI = computer-assisted personal interview; CATI = computer-assisted telephone interview; PAPI = paper-and-pencil interview.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 32 provides parent coverage rates by school and student characteristics. Approximately 85 percent of subsampled students had a corresponding parent questionnaire. Catholic schools had a weighted coverage rate of 89 percent, while other private schools and public schools had weighted coverage rates of 86 percent and 85 percent, respectively. At 90 percent, the highest coverage rate among enrollment statuses was for transfer students, while the lowest rate was 82 percent for dropout students.

Table 32. Parent coverage rates, by selected school and student characteristics: 2012

School and student characteristic	Completed students with a parent in subsample[1]	Student participants with parent coverage		
		Number	Weighted percent[2]	Unweighted percent
Total	9,754	8,296	84.8	85.1
School type				
Public	8,214	6,965	84.6	84.8
Total private	1,540	1,331	87.7	86.4
Catholic	926	804	89.0	86.8
Other private	614	527	86.4	85.8
Enrollment status[3]				
In base-year school	8,218	6,929	84.2	84.3
Transfer	1,134	1,024	89.6	90.3
Dropout	201	167	82.9	83.1
Early graduate	93	83	89.5	89.3
Homeschooled	108	93	85.6	86.1
Sex[4]				
Male	4,970	4,200	84.6	84.5
Female	4,784	4,096	85.0	85.6
Race/ethnicity[4]				
American Indian or Alaska Native	91	75	88.2	82.4
Asian[5]	1,054	836	76.0	79.3
Black, non-Hispanic	1,282	1,093	82.8	85.3
Hispanic or Latino	1,578	1,312	84.3	83.1
White, non-Hispanic	5,605	4,858	86.2	86.7
More than one race	144	122	82.4	84.7

[1] Some 47.5 percent of the student sample had a parent included in the parent subsample. Please refer to chapter 3 for more information on the sample design.

[2] Weighted percentages use the parent base weight.

[3] Enrollment status provided by the student.

[4] Sex and racial/ethnic designations determined by data provided by (in order of preference) the sample member, parents, or school, or, in rare instances, have been imputed.

[5] The Asian category also includes Native Hawaiians and other Pacific Islanders.

NOTE: All percentages are based on the number of students within the row under consideration.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

4.4.4.1 Using Field Data Collection to Target Hard-to-Reach Parent Cases

Cases with a low propensity of becoming a respondent were identified early to allow for a field intervention. A method was implemented to identify low-propensity nonrespondents after the first 6 weeks of data collection. A field intervention for cases that were less likely to respond, and thus not likely to participate in Web or CATI modes, was implemented to reduce a possible risk of nonresponse bias.

The intervention chosen for low-propensity cases was field data collection with additional field tracing. Although an experiment was not implemented, an examination of response rates is informative. Table 33 shows response rates by quartiles of the response propensity scores across the 7,847 cases that were eligible for field intervention because they were not yet completed or were otherwise finalized. The first quartile included 1,798 cases with the lowest response propensity scores (i.e., the least likely to respond), and thus were identified as target cases for field intervention. The response rate for these targeted cases was 59.1 percent.[28] This response rate slightly outperformed the cases in the second quartile (54.5 percent) of the sample ($\mathcal{X}^2$ = 8.07, p = .0045).[29] The response rates also show evidence that the model was effective in predicting response likelihood. Approximately 77 percent of cases with response propensities in the fourth quartile (i.e., likely respondents) completed a questionnaire, which is about 17 percentage points higher than cases in the other propensity quartiles.

Table 33. Response rates of targeted cases versus all other nonresponding cases at the point of model implementation

Response propensity quartiles	Number in quartile	Number responding	Percent responding
Total	7,847	4,869	62.0
First quartile[1, 2] (least likely to respond)	1,798	1,063	59.1
Second quartile	2,020	1,101	54.5
Third quartile	2,014	1,163	57.8
Fourth quartile (most likely to respond)	2,015	1,542	76.5

[1] Targeted nonrespondent cases.

[2] The first quartile cases included 2,016 parents; however, 218 of these cases were withheld from CAPI data collection. They were deemed not suitable for CAPI data collection because they had a Spanish-speaking sample member, the case had received a final code, or the sample member had already started the web interview.

NOTE: CAPI = computer-assisted personal interview.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

[28] Among the 1,798 parent cases in the first quartile, 1,314 were offered the $20 incentive of which some 616 completed a parent interview.

[29] None of the second quartile cases were targeted for field intervention as part of the nonresponse bias treatment. However, 1,288 of the second quartile cases were offered the $20 incentive of which 658 completed a parent interview.

4.4.5 Tracing and Locating

Tracing activities began prior to data collection and continued throughout the data collection period. As mentioned in section 4.3.2, tracing and locating activities prior to data collection comprised an address update from parents, a batch database tracing update for the entire sample, and school-provided address updates as part of the enrollment status update. Intensive tracing activities occurred during the data collection after all of the contacting information for a case had been exhausted. As a result of all tracing activities and contact attempts (i.e., interviewers verifying contact with sample members) throughout the data collection period, approximately 96 percent of cases were located.

4.4.6 School Staff Results

Tables 34 and 35 show response rates and coverage rates, respectively, for each of the staff questionnaires. Response rates for school administrators and school counselors at base-year schools were nearly 99 percent each. School counselor data were collected from base-year schools only. Coverage for school administrator data was increased by the collection of administrator data from transfer schools. Of the 1,822 transfer schools identified, 1,728 were contacted to complete an administrator questionnaire and 1,346 of those participated. The remaining 94 transfer school administrators were identified too late in the data collection period to initiate a questionnaire request. The addition of the transfer school administrator collection resulted in an overall administrator coverage rate of about 96 percent (unweighted), which is 9 percentage points higher than would have been achieved if only administrators from base-year schools were asked to participate.

Table 34. School administrator and counselor response rates: 2012

Questionnaire	Eligible	Participated	Unweighted percent
School administrator	2,761	2,275	82.4
Base-year schools	939	929	98.9
Transfer schools	1,822	1,346	73.9
School counselor	939	925	98.5

NOTE: All percentages are based on the number of the sample members within the row under consideration.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 35. Summary of HSLS:09 first follow-up staff questionnaire response rates at the student-level, by selected characteristics: 2012

	School administrator questionnaires							
	All schools		Base-year schools		Transfer schools		Counselor questionnaires	
Characteristic	Weighted	Unweighted	Weighted[1]	Unweighted	Weighted[1]	Unweighted	Weighted[1]	Unweighted
Total	95.2	95.8	98.9	99.2	69.3	72.4	98.7	99.1
Participated		19,509		17,646		1,863		17,631
Eligible sample		20,358		17,784		2,574		17,784
Enrollment status[2]								
In base-year school	99.0	99.3	99.0	99.3	†	†	98.7	99.1
Transfer	68.9	73.2	†	†	68.9	73.2	†	†
Dropout	88.0	86.0	94.6	97.0	73.3	58.9	99.6	99.4
Early graduate	91.3	89.5	96.6	98.0	76.9	68.3	99.7	99.3
Homeschooled	†	†	†	†	†	†	†	†
School type								
Public	93.8	94.2	98.9	99.2	68.2	72.1	98.7	99.1
Total private	97.5	97.5	99.8	99.6	83.9	75.0	99.5	99.6
Catholic	98.4	98.4	100.0	100.0	79.6	72.1	100.0	100.0
Other private	96.5	96.2	99.7	98.9	85.6	77.2	99.0	98.9
Locale								
City	92.4	94.1	98.3	98.7	63.7	67.7	98.4	99.2
Suburban	95.1	95.6	99.2	99.3	67.5	71.5	99.0	99.0
Town	97.8	97.7	100.0	100.0	76.7	77.2	100.0	100.0
Rural	96.9	96.9	98.8	99.3	77.9	77.8	98.3	98.8

See notes at end of table.

Table 35. Summary of HSLS:09 first follow-up staff questionnaire response rates at the student-level, by selected characteristics: 2012—Continued

Characteristic	School administrator questionnaires: All schools, Weighted[1]	School administrator questionnaires: All schools, Unweighted	School administrator questionnaires: Base-year schools, Weighted[1]	School administrator questionnaires: Base-year schools, Unweighted	School administrator questionnaires: Transfer schools, Weighted[1]	School administrator questionnaires: Transfer schools, Unweighted	Counselor questionnaires: Weighted[1]	Counselor questionnaires: Unweighted
Region								
Northeast	97.3	97.7	100.0	100.0	71.7	74.8	100.0	100.0
Midwest	95.6	96.6	98.6	99.5	72.4	74.7	99.2	99.0
South	95.1	95.8	99.0	99.1	68.2	72.3	98.1	98.9
West	93.4	92.9	98.3	98.2	67.5	69.0	98.2	99.0
Sex[2]								
Male	95.0	96.0	99.0	99.3	66.5	72.1	98.6	99.1
Female	95.5	95.7	98.8	99.1	72.1	72.7	98.9	99.2
Race/ethnicity[3]								
American Indian or Alaska Native	93.0	92.4	98.1	98.7	61.5	60.0	99.6	99.4
Asian[4]	97.5	97.9	99.8	99.7	72.5	81.2	99.1	99.0
Black, non-Hispanic	90.9	91.9	97.4	95.9	61.6	64.6	98.8	99.2
Hispanic or Latino	92.4	93.3	98.5	98.2	62.2	66.2	97.3	98.1
White, non-Hispanic	93.2	92.5	99.6	99.0	62.6	59.7	99.1	99.4
More than one race	97.4	97.2	99.4	99.5	78.2	77.2	99.9	99.7

† Not applicable

[1] Weighted percentages use the student base weight.

[2] Unweighted response rates are based on the number of eligible students in each column; 20,358 students linked to a base-year or transfer school administrator; 17,784 students linked to a base-year administrator or counselor; 2,574 students linked to a transfer school administrator.

[3] Sex and racial/ethnic designations determined by data provided by (in order of preference) the sample member, parents, or school, or, in rare instances, have been imputed.

[4] The Asian category also includes Native Hawaiians and other Pacific Islanders.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Chapter 5. Data Preparation and Processing

This chapter documents the automated systems, data processing, cleaning, and editing activities of the High School Longitudinal Study of 2009 (HSLS:09) first follow-up.

5.1 Overview of Systems Design, Development, and Testing

Many of the systems and processes used in the HSLS:09 first follow-up were designed during the first follow-up field test with improvements implemented for the main study. Some of the systems were adapted from the base-year study with continued improvements and adaptations for the first follow-up. The following systems were developed and used for the first follow-up:

- Integrated Management System (IMS)—a comprehensive tool used to exchange files between RTI and National Center for Education Statistics (NCES), post daily production reports, and provide access to a centralized repository of project data and documents;
- Survey Control System (SCS)—the central repository of the status of each activity for each case in the study;
- Hatteras Survey Engine and Survey Editor—a web-based application used to develop and administer the HSLS:09 instruments including the student, parent, and school staff questionnaires;
- Web-based student assessment application;
- Parent/student computer-assisted telephone interview (CATI) Case Management System (CMS)—a call scheduler and case delivery tracking system for parent telephone interviews;
- Integrated Field Management System (IFMS)—a field reporting system to help field supervisors track the status of in-school data collection and field interviewing;
- HSLS:09 survey website—public website hosted at NCES and used to disseminate information, collect sample data, and administer HSLS:09 surveys;
- School contacting system—a web-based application used to track school recruitment progress and logistics for the in-school data collection;
- Sojourn—a live CD to boot school PCs into a standard and secure environment for student in-school data collection;
- Data-cleaning programs—SAS programs developed to apply reserve code values where data are missing, clean up inconsistencies (because of respondents backing up), and fill data where answers are known from previously answered items;

- Occupation, field of study, secondary school, and postsecondary institution coding applications; and
- Data entry application for parent paper-and-pencil interview (PAPI) forms.

System development included the following phases: planning, design, development, testing, and execution and monitoring. Specifications were developed in word processing documents and flowchart applications. Specifications were updated to reflect what changed between the field test questionnaires and the full-scale questionnaires.

Each system implements safeguards to handle personally identifying information (PII) as applicable. Systems such as the IMS, Hatteras, and the SCS are standard RTI systems used successfully in the HSLS:09 base year and past studies, and were developed using the latest software tools such as Microsoft .NET and Microsoft SQL Server database.

Processing of PII by all systems was developed in accordance with the Federal Information Processing Standards (FIPS) moderate security standard. Movement of the data containing PII was handled appropriately, meeting the security requirements between the locations. Data when moved between locations were encrypted, which met the FIPS 140.2 standards, and were decrypted once they successfully reached the destination. The automated systems were developed to handle the need of moving data and files between locations in an efficient and secure manner.

5.2 Data Processing and File Preparation

Item documentation procedures were developed to capture variable and value labels for each item. Item wording for each question was also provided as part of the documentation. This information was loaded into a documentation database that could export final data file layouts and format statements used to produce formatted frequencies for review. The documentation database also had tools to produce final Electronic Codebook (ECB) input files.

5.3 Data Cleaning and Editing

Questionnaire data were stored in an SQL database that was consistent across data collection modes for a particular questionnaire. The instrument used to administer the web survey was the same instrument as the CATI, and the questionnaire data were stored in the same SQL database. This ensured that skip patterns were consistent across applications.

Editing programs were developed to identify inconsistent items across logical patterns within the questionnaire. These items were reviewed, and rules were written to either correct previously answered (or unanswered) questions to match the dependent items or blank out subsequent items to stay consistent with previously answered items.

Programs were also developed to review consistencies across multiple sources of data and identify discrepancies that required further review and resolution. Consistency checks

included unlikely patterns across rounds (i.e., between base year and first follow-up) as well as across sources within a given round (e.g., between parent and student reports).

5.4 Coding

The survey instruments collected data on major fields of study, occupations, postsecondary institutions, secondary schools, and high school math and science courses, all of which required coding. All survey instruments except the student instrument, Spanish version of the parent instrument, and the parent paper-and-pencil abbreviated questionnaire included applications which allowed respondents or interviewers to code text strings to widely used taxonomies. All text strings that were not coded during the interview were coded as part of data processing. This section describes the types of data requiring coding, the coding applications, the coding process, quality control procedures, and measures of coding quality.

5.4.1 Major Field of Study Coding

School administrators and parents identified the field of study for their most advanced postsecondary degree. If school administrators had earned a Master's degree or higher, they also reported the field of study for their Bachelor's degree. The instruments included a coding application that allowed online coding using the NCES 2010 Classification of Instructional Programs (CIP) taxonomy.[30] On the restricted-use data file, researchers will find both a 2-digit version and a 6-digit version of the CIP code for administrators' (A2HIMAJ2; A2HIMAJ6; A2BAMAJ2; A2BAMAJ6) and parents' (P2HIMAJ21; P2HIMAJ61; P2HIMAJ22; P2HIMAJ62) fields of study. Only the 2-digit versions of these variables appear on the public-use data file.

5.4.1.1 Major Field of Study Coding Methods

To use the coding application, respondents or interviewers first entered text to describe the field of study. A list of majors, customized based on the text string, was presented. The respondent or interviewer could choose one of the options listed, or choose "none of the above." If "none of the above" was selected, a two-tiered dropdown menu appeared. The first dropdown menu contained a general list of majors; the second was more specific and was dependent on the first. Interviewers were trained to use probing techniques to assist in the online coding process. Self-administered web respondents were provided supporting text on-screen. If the respondent or interviewer was unable to find a good match, he or she could proceed with the interview without selecting a code. In this case, the text string and any selections from the dropdown menus were retained.

All major text strings that were not coded during the interview were processed by RTI. First, the most commonly reported parent major text strings and all of the school administrator major text strings that appeared more than once were assigned a code by an expert coder. This

[30] For more information on this taxonomy, see http://nces.ed.gov/ipeds/cipcode/.

code was then applied to all other exact matching text strings to ensure consistency of codes for duplicate text strings. The remainder of the text strings were "upcoded" to the CIP taxonomy by coding experts using an application that used the same search function as the application in the instruments. The coding expert could assign a CIP field of study code or assign a value of 999999 to indicate that the text string was too vague to code.

5.4.1.2 Major Field of Study Coding Quality Control Procedures and Results

To evaluate the quality of the coding completed during the interview, a random sample of approximately 10 percent of the pairs of verbatim strings and codes was selected for recoding and analysis. RTI coding personnel evaluated text strings and assigned codes without knowledge of the codes that were selected during the interview. If the code selected differed from the code assigned during the interview, the coding expert was then shown both codes. The coding expert was instructed to only override the code selected during the interview if it was clearly incorrect. When a code was overridden, the new code was included on the data file in place of the original code. Text strings were designated "too vague to code" when they lacked sufficient clarity or specificity.

Results of recoding of strings coded during the interview are given in table 36. Nearly all of the codes selected during the interview were deemed to be accurate to the most detailed 6-digit level (98.1 percent). The coding expert disagreed with the CIP code selected during the interview for only 2 percent of the strings. In these instances, the expert coder's selection replaced the code selected during the interview.

Table 36. Results of quality control recoding and upcoding of major: 2012

	Percent
Sample of strings coded during interview	
Match at 6-digit and 2-digit	98.1
Match at 2-digit, but not 6-digit	0.0
Disagree	1.9
Sample of strings coded during data processing	
Match at 6-digit and 2-digit	65.2
Match at 2-digit, but not 6-digit	17.3
Match at too vague to code	5.7
Disagree	11.8 [1]

[1] All of the majors that were not coded during the interview were coded independently by two coding experts and compared. This figure includes instances where one coding expert thought the string was too vague to code and the other did not. A third coding expert adjudicated all of the instances of disagreement.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

The coding of major text strings by RTI's expert coders was also subject to a quality control review. Strings that were not coded during the interview and not batch coded were

upcoded to the CIP taxonomy by a coding expert. All of these upcodes were selected for independent coding by a second coding expert. When this process was complete, the results of the two coding experts were compared. The results are shown in table 36.

The two expert coders selected the same detailed 6-digit code for 65 percent of the text strings. The expert coders disagreed at the 6-digit level, but agreed at the 2-digit level for 17 percent of the text strings. Combining these with the strings where agreement was reached at both levels yields 83 percent agreement at the 2-digit level overall. Both of the expert coders determined that the text was too vague to code for 6 percent of the strings. Disagreement occurred for 12 percent of the strings.

All instances where there was disagreement at either the 6-digit or 2-digit level or where one expert coder thought the string was too vague to code were adjudicated by a third expert coder who had done the batch coding.

5.4.2 Parent Occupation Coding

The HSLS:09 first follow-up parent instrument included tools that allowed online coding of literal responses of occupation job title and duties to the 2000 Standard Occupational Classification(SOC) taxonomy.[31] The 2000 SOC taxonomy was used for consistency with the base-year data. Occupation job title and duties were matched to occupation descriptions from the Occupational Information Network (O*NET).[32] Respondents to the parent interview were asked to identify the occupation of up to two parents or guardians in the household, typically their own occupation and the occupation of a spouse or partner. Additionally, students who did not have a parent complete the parent survey in the base year were asked to provide the occupation for up to two parents in their household. For technical information on these variables, see appendix X.

On the restricted-use data file, researchers will find both a 2-digit version and a 6-digit version of the O*NET code for parents' occupations:

- X2PAR1OCC2 [X2 Parent 1: Current/most recent occupation: 2-digit O*NET code],
- X2PAR1OCC6 [X2 Parent 1: Current/most recent occupation: 6-digit O*NET code],
- X2PAR2OCC2 [X2 Parent 2: Current/most recent occupation: 2-digit O*NET code],
- X2PAR2OCC6 [X2 Parent 2: Current/most recent occupation: 6-digit O*NET code],
- X2MOMOCC2 [X2 Mother/female guardian's current/most recent occupation: 2-digit O*NET code],
- X2MOMOCC6 [X2 Mother/female guardian's current/most recent occupation: 6-digit O*NET code],
- X2DADOCC2 [X2 Father/male guardian's current/most recent occupation: 2-digit O*NET code], and

[31] See http://www.bls.gov/soc/2000/socstruc.pdf.
[32] See http://www.onetcenter.org/overview.html.

- X2DADOCC6 [X2 Father/male guardian's current/most recent occupation: 6-digit O*NET code].

Only the 2-digit versions of these variables appear on the public-use data file.

5.4.2.1 Parent Occupation Coding Methods

Respondents were asked to provide a job title and job duties for each occupation. All parents completing the English version of the web interview were also asked to code the occupation to the SOC taxonomy. To code the occupation, interviewers or respondents who were self-administering the web questionnaire first entered the job title and duties. These strings were automatically matched to the occupation descriptions from O*NET and a customized list of occupations was presented. Interviewers and respondents could choose one of the options listed, or choose "none of the above." In the occupation coding application, selecting "none of the above" presented the user with a set of three sequential dropdown menus, each with choices increasing in their level of specificity. The first dropdown menu contained a general list of occupations. The options presented in the second dropdown were dependent on the code selected in the first. Some selections from the second dropdown required users to make a selection from a third even more detailed dropdown menu. Interviewers were trained to use probing techniques to assist in the online coding process. Self-administered web respondents were provided supporting text on screen. If the respondent or interviewer was unable to find an appropriate SOC code for the occupation, they could proceed with the interview without selecting a code. In this case, the text string and any selections from the dropdown menus were retained to assist with coding during data processing.

RTI's coding experts attempted to code all occupations that were not coded in the web interview or on the PAPI. This "upcoding" was completed using an application that used the same search function as the application in the parent instrument. The coding expert could assign an O*NET code or assign a value of 999999 to indicate that the text string was too vague to code.

5.4.2.2 Parent Occupation Coding Quality Control Procedures and Results

Coding experts evaluated the quality of coding that was completed during the interview by recoding a random sample of approximately 20 percent of the occupation text strings. To recode the selected occupations, RTI staff members worked with a coding application which used the same search function as the application in the instruments. These coding experts evaluated text strings and assigned codes without knowledge of the codes that were selected during the interview. If the code selected differed from the code assigned during the interview, the coding expert was then shown both codes. The coding expert was instructed to only override the code selected during the interview if it was clearly incorrect. When the expert coder did not agree with the 6-digit code selected during the interview, the new code was included on the data

file in place of the original code. Text strings were designated "too vague to code" when they lacked sufficient clarity or specificity.

The expert coders agreed with the 6-digit code selected during the interview for 84 percent of the text strings reviewed and agreed with the 2-digit code (but not the 6-digit code) for an additional 7 percent of the text strings reviewed, for a total of 91 percent agreement at the 2-digit level. The expert coder disagreed with the code selected during the interview for 8 percent of the occupations.

The coding of parent occupation text strings by RTI's expert coders was also subject to a quality control review. All of the text strings that were not coded during the interview were upcoded to the O*NET taxonomy by a coding expert. Approximately 20 percent of these upcodes were randomly selected for independent coding by a second coding expert. When this process was complete, the results of the two coding experts were compared. The results are shown in table 37.

Table 37. Results of quality control recoding and upcoding of parent occupation coding: 2012

	Percent
Sample of strings coded during interview	
Match at 6-digit and 2-digit	84.2
Match at 2-digit, but not 6-digit	7.0
Disagree	8.8
Sample of strings coded during data processing	
Match at 6-digit and 2-digit	78.9
Match at 2-digit, but not 6-digit	9.7
Match at too vague to code	2.3
Disagree	9.2 [1]

[1] This figure includes instances where one coding expert thought the occupation was too vague to code and the other did not.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

The two expert coders selected the same detailed 6-digit code for 79 percent of the text strings. The coders agreed at the 2-digit level but not the 6-digit level for 10 percent of the strings. Summing these two categories yields 89 percent in agreement at the 2-digit level. A very small percentage of strings were deemed too vague to code by both expert coders (2 percent). The coders disagreed for 9 percent of the strings. If the two coders disagreed, the second expert coder's selection appears on the data file.

5.4.3 Student Job at Age 30 Coding

The HSLS:09 first follow-up student instrument asked respondents to indicate what occupation they thought they would have when they were age 30. Students entered a job title, but were not asked to enter job duties. Respondents also had the option of checking a box to indicate

that they did not know. On the restricted-use data file, researchers will find both a 2-digit version and a more specific 6-digit version of the 2000 O*NET code for students' job at age 30:

- X2STU30OCC2 [X2 Student occupation at age 30: 2-digit O*NET code], and
- X2STU30OCC6 [X2 Student occupation at age 30: 6-digit O*NET code].

Only the 2-digit version of this variable appears on the public-use data file. For technical information on these variables, see appendix E (Composite Variables).

5.4.3.1 Student Job at Age 30 Coding Methods

Students were not asked to code their expected occupations so all job titles needed to be coded after data collection using the O*NET taxonomy. The text strings were first matched against coded strings from the base year. When text strings matched between base year and first follow-up, the base-year code was applied to the first follow-up text string. As shown in table 38, of 14,666 text strings provided by students, 9,226 strings (63 percent) were coded using base-year codes. The 5,440 text strings that remained uncoded by this method were worked by expert coders using the same lookup application that was used in the parent instrument to code parent occupations. However, because job duties were not collected for students' expected jobs, the search for an appropriate O*NET code was strictly based on the job title.

Table 38. Job at age 30 text strings, by coding method: 2012

		Matched to base-year text string		Coded by expert coders	
	Total text strings	Number	Percent	Number	Percent
Job at age 30 text strings	14,666	9,226	62.9	5,440	37.1

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

5.4.3.2 Student Job at Age 30 Quality Control Procedures and Results

For quality control purposes, a subset (10 percent) of the initially coded cases were randomly selected, and then independently coded by a second coding expert. When the second result was different from the first result, both results were displayed. The second coder could then agree with the first coder or override the first coder's result. When this process was completed, the results of the two coding experts were compared. As shown in table 39, the two coding experts arrived at the same results (a match at the 6- and 2-digit level) for 62 percent of the text strings, and agreed on the same 2-digit code (but not the 6-digit level) for 18 percent. Both coders determined the string was too vague to code for 1 percent of the cases. Disagreement occurred in 19 percent of the cases. It may be noted that the rate of disagreement for students' expected jobs is higher than the rate of disagreement for parent occupations (19 percent versus 9 percent). This was expected given that students were reporting hypothetical jobs and therefore were not asked to describe job duties. When coders disagreed, a third coding expert adjudicated.

Table 39. Results of quality control of job at age 30 upcoding: 2012

Results	Percent
Total coded by second coding expert	100.0
Match at 6-digit and 2-digit	61.5
Match at 2-digit, but not 6-digit	17.9
Match at too vague to code	1.3
Disagree	19.3

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

After all coding was completed, about 99 percent of the text strings had been assigned a 6-digit O*NET code, the most specific level; 1 percent were coded to a lesser level of specificity (2-digit level). Only 1 percent were too vague to code at all.

5.4.4 High School Coding

Teenagers were asked to provide the name, city, and state of all high schools they had attended other than the school they were attending in the base year. These included schools the respondent attended at the time of the first follow-up data collection (S2HSID), schools the respondent last attended before leaving or completing high school (S2LASTHSID), and schools attended between enrollment at the base-year high school and the most recent high school (S2OTHHSID1-4).

5.4.4.1 High School Coding Methods

The survey contractor matched the text strings to two NCES databases: (1) the Common Core of Data (CCD), a comprehensive and annually updated database on public elementary and secondary schools and school districts in the United States; and (2) the Private School Universe Survey (PSS), a similar database of private schools in the United States. Multiple years of each database were searched to account for schools that had closed and opened over the time span of the HSLS:09 base-year and first-follow-up surveys. Web searches were conducted to locate schools that were not found in the CCD or PSS. These tended to be schools that had opened recently, alternative schools, and charter schools.

5.4.4.2 High School Coding Results

The results of coding by variable are presented in table 40. The majority of schools named by students were located in the United States and will be contacted for transcript collection; 96.7 percent of transfer schools attended at the time of the first follow-up interview, 90.9 percent of the schools that were last attended by those who were not enrolled in school, and 87.6 percent of all other high schools named.

Table 40. Final disposition of high school coding: 2012

	Number	Percent
Current transfer school		
Located in United States	2,299	96.7
Unlocated or foreign	78	3.3
Last school attended		
Located in United States	179	90.9
Unlocated or foreign	18	9.1
Other high schools		
Located in United States	1,022	87.6
Unlocated or foreign	145	12.4

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

5.4.5 Postsecondary Institution (Integrated Postsecondary Education Data System [IPEDS]) Coding

Parents and students were each asked to indicate their first choice of postsecondary institutions for the sample member to attend, and the postsecondary institution the sample member would most likely attend. The postsecondary institution IDs are available only on the restricted-use data file:

- S2LIKELYCLGID [IPEDS ID of postsecondary institution teen most likely to attend in 2013],
- S2CHOICECLGID [IPEDS ID of teen's first choice postsecondary institution in 2013],
- P2LIKELYCLGID [IPEDS ID of postsecondary institution parent thinks teen most likely to attend in 2013], and
- P2CHOICECLGID [IPEDS ID of parent's first choice postsecondary institution for teen in 2013].

5.4.5.1 Postsecondary Institution Coding Methods and Results

Although students were not asked to code their institution, parents were asked to indicate the postsecondary institutions using a look-up tool. After parents (or the interviewer) entered the institution's name, city, and state into the web survey, they could search an online look-up tool containing institutions from the 2010–11 IPEDS for the appropriate match. When a match was not found, the respondent was asked to provide the institution's level (i.e., 4-year, 2-year, less-than-2-year) and control (i.e., public, private not-for-profit, private-for-profit). This information was later used to assist RTI staff in finding a match in IPEDS as part of data processing.

Text strings provided by students and text strings not coded by parents through the online look-up tool were provided to coding experts to be upcoded in the following manner. First, cases

were compared against the 2010–11 IPEDS database for matching. Any case with school name, city, and state that exactly matched an IPEDS record was assigned the corresponding IPEDS ID. Then, any text strings that remained uncoded were loaded into the coding application for an RTI coding expert to assign IDs. As shown in table 41, 18 percent of all text strings were coded during the interview and 82 percent were coded by expert coders at RTI.

Table 41. Postsecondary institution names, by IPEDS coding method: 2012

		Coded during interview		Coded by expert coders	
Respondent	Total	Number	Percent	Number	Percent
Total	18,484	3,357	18.2	15,127	81.8
Student provided	14,122	0	0.0	14,122	100.0
Parent provided	4,362	3,357	77.0	1,005	23.0

NOTE: Detail may not sum to totals because of rounding. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

After all coding was complete, 99 percent of the institutions had been assigned an IPEDS ID. Less than 1 percent of institutions were foreign, and 1 percent were uncodeable, which usually meant the string was not a postsecondary institution.

5.4.6 Math and Science Course Coding

In the student instrument, students were asked to select the math and science, engineering, and computer science courses that they were taking as of the spring semester of 2012. Students were able to select more than one course. They were also able to select "other" and report the course title. The "other" courses were upcoded to the existing course variables by a math and science curriculum expert.

Math and Science Coding Methods and Results

All course titles were coded in a Microsoft Excel spreadsheet which allowed for sorting to ensure that all similar course titles would be coded consistently. The results for math course coding are presented in table 42. There were 836 math courses listed in the "other specify" text field. Nearly one half of the course titles were upcoded to either "Business, Consumer, General, Applied, Technical, Functional, or Review Math" (24 percent) or "Other Statistics and Probability" (23 percent). New categories were created for "College Algebra/College Math," "Advanced/Accelerated Math," "Liberal Arts Math," and "Math Topics" given their prevalence and the lack of an appropriate existing category.

Table 42. Distribution of math course upcodes: 2012

	Number	Percent
Total	836	100.0
Pre-Algebra	5	0.6
Algebra I, 1A, or 1B	40	4.8
Algebra II	37	4.4
Algebra III	4	0.5
Geometry	8	1.0
Trigonometry	23	2.8
Pre-calculus or Analysis and Functions	45	5.4
Other Calculus	2	0.2
Other Statistics or Probability	192	23.0
Integrated Math I	5	0.6
Integrated Math II	2	0.2
Integrated Math III or above	4	0.5
International Baccalaureate (IB) mathematics standard level	3	0.4
Business, Consumer, General, Applied, Technical, Functional, or Review Math	199	23.8
Other—College Algebra/College Math	95	11.4
Other—Advanced/Accelerated Math	40	4.8
Other—Liberal Arts Math	20	2.4
Other—Math topics	15	1.8
Other math course	97	11.6

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

The results for science, engineering, and computer science course coding are presented in table 43. There were 1,108 science, engineering, or computer science courses listed in the other specify text field. Over one-third of these courses were able to be upcoded to “Other biological sciences such as botany, marine biology, or zoology.” Almost half of these other biological science courses were in forensic science. The “Other earth or environmental sciences such as ecology, geology, oceanography, or meteorology” and “Computer applications” categories each accounted for about another 10 percent of the course titles.

Table 43. Distribution of science course upcodes: 2012

	Number	Percent
Total	1,108	100.0
Life science	2	0.2
Biology I	9	0.8
Biology II	4	0.4
Anatomy or Physiology	42	3.8
Other biological sciences such as botany, marine biology, or zoology	410	37.0
Chemistry I	44	4.0
Chemistry II	20	1.8
Earth Science	7	0.6
Advanced Placement (AP) Environmental Science	3	0.3
International Baccalaureate (IB) Environmental Systems and Societies	1	0.1
Other earth or environmental sciences such as ecology, geology, oceanography, or meteorology	107	9.7
Physics I	14	1.3
Physics II	5	0.5
Physical Science	1	0.1
Principles of Technology	10	0.9
Other physical sciences such as astronomy or electronics	48	4.3
Integrated Science I	21	1.9
General Science	27	2.4
Computer Applications	108	9.7
Computer Programming	12	1.1
Advanced Placement (AP) Computer Science	1	0.1
International Baccalaureate (IB) Design Technology	2	0.2
Other computer or information science course	37	3.3
An engineering course such as general engineering, robotics, aeronautical, mechanical or electrical engineering	38	3.4
Other science, computer science, or engineering course	135	12.2

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

5.5 Construction of Scales

Certain sets of items that appear in the student and administrator surveys were designed to be analyzed as psychological scales. Scales need to be unidimensional in terms of their factor structure, and reliable, and indeed, increase the reliability above that that could be obtained through use of a single item for measurement purposes. The student survey includes, for example, questions related to science and mathematics identity and self-efficacy, while the school administrator survey includes scale items pertaining to school problems and ethos or climate. Most of the scales included in this dataset are conceptually similar to scales developed in previous education studies such as ELS:2002; for example, math and science self-efficacy,

utility, and effort. Development of these scales was also based on advice received from HSLS:09 Technical Review Panel (TRP) members and TRP meeting participants.

Prior to constructing the scales, questionnaire responses were subjected to data cleaning procedures discussed previously. Questionnaire items were reverse coded (i.e., positively and negatively worded items were coded to reflect the same direction on the construct) to equate larger scale values with positive attributes (e.g., higher levels of self-efficacy). Once the data were finalized, the (weighted) reliability of the scale items was evaluated using Cronbach's alpha. Weighted scales were then created if the associated items had at least a 0.65 alpha level with SAS® *Proc Factor* and standardized to have mean zero and (weighted) standard deviation of one. (All scales exceeded the 0.65 alpha criterion). Scales were set to missing if any of the scale items were missing. The individual item-level data are also available on the data file. Researchers are encouraged to further examine the psychometric properties of the scales using the item-level data. The scales presented on the data file are just one way to combine the information. Individual researchers are able to recreate these or similar scales.

Eleven scales were created at the student level. These included a mathematics identity scale (X2MTHID); mathematics utility scale (X2MTHUTI); mathematics self-efficacy scale (X2MTHEFF); mathematics interest scale (X2MTHINT); mathematics effort scale (X2MEFFORT); science identity scale (X2SCIID); science utility scale (X2SCIUTI); science self-efficacy scale (X2SCIEFF); science interest scale (X2SCIINT); science effort scale (X2SEFFORT); and student behavior scale (X2BEHAVEIN). Two base-year scales were not re-created for the first follow-up data collection: school belonging (X1SCHOOLBEL), and school engagement, (X1SCHOOLENG). Three new scales are included in the first follow-up dataset, including math effort (X2MEFFORT), science effort (X2SEFFORT), and student behavior (X2BEHAVEIN). The input variables used for the first follow-up are the same as those used for the base-year variables, except X2MTHINT and X2SCIINT, which do not include a first follow-up analogue to S1LEASTSUBJ because the first follow-up instrument did not ask for students to identify their least favorite subject. The student-level scales are included in table 44.

Two scales were created at the school level. These include a school climate scale (X2SCHOOLCLI) and a school problems scale (X2PROBLEM). The school climate scale (X2SCHOOLCLI) includes 13 of the 14 variables from the base-year version, X1SCHOOLCLI. Instead of A1BULLY, this scale includes two additional input variables, A2CYBERBULLY and A2OTHERBULLY. X2PROBLEM is a new scale included in the first follow-up data set. The school-level scales are specified in table 45.

Table 44. Summary information for student scales: 2012

Student scale	Variable name	Cronbach's alpha
X2MTHID: Mathematics Identity	S2MPERSON1	0.88
	S2MPERSON2	
X2MTHUTI: Mathematics Utility	S2MUSELIFE	0.82
	S2MUSECLG	
	S2MUSEJOB	
X2MTHEFF: Mathematics Efficacy	S2MTESTS	0.89
	S2MTEXTBOOK	
	S2MSKILLS	
	S2MASSEXCL	
X2MTHINT: Mathematics Interest	S2MWASTE	0.69
	S2MBORING	
	S2FAVSUBJ	
	S2MENJOYS	
	S2MENJOYING	
X2MEFFORT: Mathematics Effort	S2MATTENTION	0.74
	S2MONTIME	
	S2MSTOPTRYING	
	S2MGETBY	
X2SCIID: Science Identity	S2SPERSON1	0.89
	S2SPERSON2	
X2SCIUTI: Science Utility	S2SUSELIFE	0.82
	S2SUSECLG	
	S2SUSEJOB	
X2SCIEFF: Science Efficacy	S2STESTS	0.92
	S2STEXTBOOK	
	S2SSKILLS	
	S2SASSEXCL	
X2SCIINT: Science Interest	S2SWASTE	0.77
	S2SBORING	
	S2FAVSUBJ	
	S2SENJOYS	
	S2SENJOYING	
X2EFFORT: Science Effort	S2SATTENTION	0.76
	S2SONTIME	
	S2SSTOPTRYING	
	S2SGETBY	

See notes at end of table.

Table 44. Summary information for student scales: 2012 (continued)

Student scale	Variable name	Cronbach's alpha
X2BEHAVEIN: Student Behavior	S2LATESCH	0.73
	S2SKIPCLASS	
	S2INSCHSUSP	
	S2ABSENT	
	S2WOHWDN	
	S2WOPAPER	
	S2WOBOOKS	

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 45. Summary information for school scales: 2012

School scale	Variable names	Cronbach's alpha
X2SCHOOLCLI: School Climate	A2CONFLICT	0.88
	A2ROBBERY	
	A2VANDALISM	
	A2DRUGUSE	
	A2ALCOHOL	
	A2DRUGSALE	
	A2WEAPONS	
	A2PHYSABUSE	
	A2TENSION	
	A2VERBAL	
	A2MISBEHAVE	
	A2DISRESPECT	
	A2GANG	
	A2CYBERBULLY	
	A2OTHERBULLY	
X2PROBLEM: School Problems	A2TARDY	0.86
	A2STUABSENT	
	A2CUT	
	A2DROPOUT	
	A2APATHY	
	A2PRNTINV	
	A2RESOURCES	
	A2UNPREP	
	A2HEALTH	

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Chapter 6. Analytic Weights, Variance Estimation, and Nonresponse Bias Analysis

6.1 Overview of Weighting in Base Year and First Follow-up

Estimates are generated for the High School Longitudinal Study of 2009 (HSLS:09) target populations with a set of analytic weights and software that accounts for the complex, two-stage sample design. A series of weights have been computed for HSLS:09 to accommodate analyses specific to each round of the study (base year or first follow-up) plus analyses to evaluate change across a 2- to 3-year time period.

Five sets of weights were constructed for the HSLS:09 base year: a school-level weight used to analyze information collected in the administrator and counselor questionnaires; a student weight for analyzing student survey responses and mathematics ability; and three contextual weights to incorporate responses obtained from the science teacher questionnaire, the mathematics teacher questionnaire, and the home-life (parent) questionnaires. The steps implemented to create these weights are detailed in the HSLS:09 Base-year Data File Documentation (Ingels et al. 2011). Relevant information as it pertains to the first follow-up is provided below to strengthen the discussion.

Four analytic weights were computed for the HSLS:09 first follow-up using a similar methodology as implemented in the base year. They include two student weights, one for analyses specific to the first follow-up (section 6.3.1) and one for longitudinal analyses associated with change between the base year and first follow-up (section 6.3.2), and two home-life contextual weights, one weight for first follow-up analyses (section 6.4.1) and one for longitudinal analyses (section 6.4.2), connected to responses obtained from the parent questionnaire. Base weights and weight adjustments applied to construct the final analytic weights are discussed in each section, as well as steps implemented to construct the corresponding sets of balanced repeated replication (BRR) weights for variance estimation. Each section follows from the discussion in section 6.2, which identifies the appropriate HSLS:09 analytic weight for a set of example analyses.

Precision (standard errors) and bias are important attributes to evaluate when assessing the quality of the survey estimates. Precision for a set of important characteristics is summarized in section 6.5, following specifications for correctly calculating HSLS:09 standard errors (details are provided in appendix F). Results from a series of nonresponse bias analyses are summarized in section 6.6 to highlight the effectiveness of the weight adjustments in improving data quality (details are provided in appendix G). Complementing the discussion of quality in the estimates,

section 6.7 concludes this chapter with a description of the quality control procedures used to construct the weights.

6.2 Choosing an Analytic Weight

The choice and number of HSLS:09 weights were driven by the need to maximize the analytic utility for the research community. Such analyses may include responses obtained from a particular instrument within a round of the study (e.g., student questionnaire responses in the base year) along with combinations of data from multiple instruments (e.g., student and parent questionnaire responses, change in student responses from base year to first follow-up). As discussed in the base-year documentation and repeated here, weights were derived to facilitate many but not all possible scenarios.[33] Some important scenarios are described below.

The first follow-up data file contains a total of nine analytic weights (table 46): five weights for analysis of the base-year data and four weights to be used in conjunction with the first follow-up data (two weights for analysis of first follow-up responses, and two weights for analysis of population change from base year to first follow-up[34]). In summary, researchers analyzing *any* data from the first follow-up (alone or in conjunction with base-year data) should use one of the four first follow-up weights. Analyses involving *only* the base-year data, with no first follow-up data, should include one of the five weights for analysis of base-year data.

The base-year school and student-level weights contain adjustments for base-year nonresponse of schools. The base-year student-level weights additionally contain adjustments for nonresponse patterns among students and, for the contextual weights, nonresponse patterns linked to parents or teachers. The first follow-up weights were adjusted for school-level nonresponse in the base year and either student nonresponse (W2STUDENT) or parent nonresponse (W2PARENT) in the HSLS:09 first follow-up. The two longitudinal weights, described in the last rows of table 46, were adjusted for school-level nonresponse in the base year and either student nonresponse in the base year and first follow-up (W2W1STU), or student and parent nonresponse in the base year and first follow-up (W2W1STU). Response patterns for the HSLS:09 base year and first follow-up addressed by the weights are summarized in table 47.

The following guidelines are provided to assist researchers in identifying the appropriate weight for analyses that include a particular combination of components. The base-year information below has been included from the base-year documentation for completeness.

[33] The creation of additional HSLS:09 weights was considered to address other analytic scenarios. However, to limit the size of the analytic files and to limit potential confusion in the choice of analytic weight if a large number of weights were produced, decisions were made to focus only on the most likely types of analyses given the HSLS:09 first follow-up data sources.

[34] Population change is also referred to as "gross change."

Table 46. HSLS:09 analytic weights

HSLS:09 study	Universe[1]	Estimation	Variable name	Nonresponse-adjusted component(s) in each weight[2]
Base year	All study-eligible schools	Base year only	W1SCHOOL	School
Base year	All study-eligible students in base year[3]	Base year only	W1STUDENT W1PARENT W1SCITCH W1MATHTCH	Student Student*Parent Student*Science teacher Student*Math teacher
First follow-up	9th-grade cohort	First follow-up only	W2STUDENT W2PARENT	Student Parent
Base year and first follow-up	9th-grade cohort	Change from base year to first follow-up	W2W1STU W2W1PAR	Student[4] Student*Parent[4]

[1] The sum of the associated analytic weights estimates the total for the universe.
[2] Student-level weights are a function of the school analytic weights and therefore are also adjusted for school nonresponse. Unless otherwise specified, the weights were additionally adjusted for nonresponse within the listed HSLS:09 study.
[3] The subpopulation associated with the public-use file is restricted to ninth-grade students who were capable of participating in the student questionnaire and math assessment.
[4] The student longitudinal weights account for nonresponse in the base year, the first follow-up, or both.
NOTE: The symbol "*" should be interpreted as "and." For example, the W1PARENT weight was developed using adjustments for student and parent nonresponse.
SOURCE: U.S. Department of Education, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Data File.

Table 47. Number and percent of completed surveys for the student sample by HSLS:09 study and instrument: 2012

HSLS:09 study	Instrument	Eligible	Participated	Weighted percent[1]	Unweighted percent
Base year	Student questionnaire	25,206	21,444	85.7	85.1
	Student assessment	25,206	20,781	83.0	82.4
	Parent questionnaire	21,444	16,429	76.1	76.6
	School administrator	21,444	20,301	94.2	94.7
	School counselor	21,444	19,505	90.2	91.0
	Teacher questionnaire				
	Math teacher	20,970	16,035	72.3	76.5
	Science teacher	20,101	14,629	70.0	72.8
First follow-up	Student questionnaire	25,184	20,594	82.0	81.8
	Student assessment	25,184	18,507	73.0	73.5
	Parent questionnaire[2]	11,952	8,651	72.5	72.4
Base year and first follow-up[3]	Student questionnaire	25,184	18,623	74.3	74.0
	Student assessment	25,184	16,356	64.7	65.0
	Parent questionnaire[4]	10,210	6,611	64.2	64.8

[1] All weighted percentages are based on the row under consideration and are calculated with the student base weight.
[2] Details of the parent subsample design are provided in section 3.3.4.
[3] Only sampled students who participated in both the base year and first follow-up are considered as participants.
[4] Participants are identified as sampled students who participated in both the base year and first follow-up and who have parent responses in both the base year and first follow-up.
SOURCE: U.S. Department of Education, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-Use Data File.

6.2.1 Base-year School-level Analysis

School level analysis is only appropriate with the base-year school-level data. At the base year, the HSLS:09 study design supports national estimates of schools with ninth-graders in the 2009-10 school year.[35]

- W1SCHOOL. This weight accounts for base-year school nonresponse. Estimates generated with this base-year school weight are associated with the HSLS:09 target population of schools. This weight can be used for the analysis of school characteristics, school administrator survey data, and counselor survey data, individually or in combination. Note that weighted values generated from the school administrator and counselor response provide information for the HSLS:09 target population of schools.[36] Additional information on construction of the school weight is provided in section 6.3.1 and in the HSLS:09 base-year documentation and is not repeated here.

6.2.2 Base-year Student-level Analysis

As mentioned, if a researcher is *only* using base-year student-level data, with no first follow-up data, then one of the four weights discussed in this section should be used. If a researcher is analyzing any data from the first follow-up (alone or in conjunction with base year data), one of the four first follow-up weights should be used (see section 6.2.3).

- W1STUDENT. This weight accounts for (1) base-year school nonresponse and (2) student nonresponse in the base-year study. All records for sample students who participated in the base year will have a positive (nonzero) weight. Estimates generated with this base-year student weight are associated with the HSLS:09 target population of ninth-grade students.[37] This weight can be used for the analysis of base-year student assessment scores or survey data, alone or in combination with the school characteristics or administrator/counselor data.
- W1PARENT. This weight accounts for nonresponse in the base year associated with (1) schools, (2) students, and (3) parents.[38] All records for sample students who participated in the base year with a parent who also participated in the base year will have a positive (nonzero) weight. Estimates generated with this base-year student home-life weight are associated with the HSLS:09 target population of ninth-grade students. This weight can be used for the analysis of base-year parent responses, alone or in conjunction with student data, school characteristics, or administrator/counselor data.

---

[35] Base-year school-level estimates pertain to all regular public schools, including public charter schools, and all private schools in the 50 United States and the District of Columbia providing instruction to students in both the 9th and 11th grades. Additional details are found in section 3.2.1 of the base year documentation.

[36] Questionnaire responses were requested from the lead counselor or counselor most knowledgeable about ninth-grade counseling practices at each sampled school. Because the counselor was *not* randomly selected from the set of counselors, contextual estimates can only be generalized to the target population of schools and not to a population of school counselors.

[37] An analysis of the nonresponse patterns in the combined student and administrator or counselor data did not indicate the need for additional student-level weights.

[38] Parent information was available for neither *all* sampled ninth-grade students nor for the target population of parents. Therefore, the contextual weights were adjusted for the known characteristics of the participating students.

- W1SCITCH. This weight includes adjustments for (1) school nonresponse, (2) student nonresponse, and (3) science-teacher nonresponse in the base year.[39] All records for sample students who participated in the base year with a science teacher who also participated in the base year will have a positive (nonzero) weight. Estimates generated with this base-year science-course enrollee weight are associated with the subgroup of ninth-grade students in the HSLS:09 target population taking a science course in the ninth grade. These estimates ***do not*** reflect the population of all science teachers of ninth-grade students because science teachers themselves were not sampled directly (see section 6.5 of the base year documentation for further information). This weight can be used for the analysis of science teacher data in conjunction with base-year student data, school characteristics, or administrator/counselor data.
- W1MATHTCH. This weight includes adjustments for (1) school nonresponse, (2) student nonresponse, and (3) mathematics-teacher nonresponse in the base year.[40] All records for sample students who participated in the base year with a mathematics teacher who also participated in the base year will have a positive (nonzero) weight. Estimates generated with this base-year mathematics-course enrollee weight are associated with the subgroup of ninth-grade students in the HSLS:09 target population taking a mathematics course in the ninth grade. These estimates ***do not*** reflect the population of all mathematics teachers of ninth-grade students because mathematics teachers themselves were not sampled directly (see section 6.5 of the base year documentation for further information). This weight can be used for the analysis that draws on mathematics teacher data in conjunction with base-year student data, school characteristics, or administrator/counselor data.

6.2.3 First Follow-up Student-level Analysis

If a researcher is analyzing any data from the first follow-up (alone or in conjunction with base year data), one of the four first follow-up weights discussed in this section should be used.

- W2STUDENT. This weight accounts for (1) base-year school nonresponse and (2) student nonresponse in the first follow-up only (regardless of the student's base-year response status). All records for sample students who participated in the first follow-up will have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of ninth-grade students.[41]

[39] As with the home-life contextual weight (W1PARENT), the science teacher contextual weights were adjusted for the known characteristics of the participating students because information was not available for the associated target population of teachers. The sum of the weights estimates the total number of ninth-grade students in the HSLS:09 target population taking a science course and is less than the total number of ninth-grade students.

[40] The mathematics teacher contextual weights were adjusted for the known characteristics of the participating students because information was not available for the associated target population of teachers. The sum of the weights estimates the total number of ninth-grade students in the HSLS:09 target population taking a mathematics course and is less than the total number of ninth-grade students.

[41] Responses in the first follow-up were obtained from the administrator and counselor of the base-year sample school for (1) students who were attending that school during the first follow-up, and (2) dropouts and early graduates whose last known school was that base-year school. First follow-up administrator responses, but not counselor responses, were obtained from the transfer school for (1) students who were attending the transfer school during the first follow-up, and (2) dropouts and early graduates who had last attended that school. Administrator and counselor responses were not obtained for home-schooled students and nonresponding transfer students.

This weight can be used for the analysis of first follow-up student assessment scores or survey data, alone or in combination with the school characteristics, administrator/counselor data from either round of HSLS:09, or teacher data from the base year.[42] If the analysis includes base year student data, researchers are encouraged to consider W2W1STU (see below).

- W2PARENT. This weight accounts for (1) base-year school nonresponse, (2) subsampling of parents for the first follow-up, and (3) parent nonresponse in the first follow-up.[43,44] All records for sample students with a parent who participated in the first follow-up will have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of ninth-grade students. This weight can be used for analysis of first follow-up parent responses alone or in combination with student survey and/or assessment data, administrator/counselor data from either round of HSLS:09, or teacher data from the base year. If the analysis includes base year parent interview data, researchers are encouraged to consider W2W1PAR (see below).

Two sources of contextual information for analysis of the student data were obtained in the HSLS:09 base year but not in the first follow-up. They include interviews with the science teacher and mathematics teacher for students taking the associated course in the ninth grade. Researchers may choose to condition the analyses of first follow-up student data on teacher responses obtained in the base year. Unlike the base-year data file, the HSLS:09 first follow-up data file does not contain contextual analytic weights to account for nonresponse among students with base-year teacher information. Instead, as discussed previously, either W2STUDENT or W2PARENT should be used depending on the inclusion of parent responses. Note that estimates generated with student data and either W2STUDENT or W2PARENT in conjunction with the base-year teacher responses are no longer associated with the HSLS:09 target population of ninth-grade students and should be used with caution.

6.2.4 Student-level Longitudinal Analysis

- W2W1STU. This weight accounts for (1) base-year school nonresponse, and (2) student nonresponse in both the base year and the first follow-up. All records for sample students who participated in the base year *and* first follow-up will have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of ninth-grade students. This weight can be used for analysis of population change that examines the student data from the base year to

42 As discussed for the course enrollee weights, not all students were taking science or mathematics courses in the ninth grade. Therefore, analyses involving the base-year teacher responses will provide estimates for the subgroup of ninth-grade students in the HSLS:09 target population taking the associated course.

43 Note, W2PARENT differs slightly from the base-year weight (W1PARENT). Unlike the base year, a positive weight was calculated for student cases with a responding parent irrespective of the student's first follow-up response status. The base-year weight was calculated only for participating students with a responding parent.

44 Parent information was available for neither *all* sampled ninth-grade students nor for the target population of parents. Therefore, the contextual weights were adjusted for the known characteristics of the participating students. Note: student data are not available for 355 responding parent records because of student nonresponse in the first follow-up.

the first follow-up, alone or in combination with administrator/counselor data from either round of HSLS:09, or teacher data from the base year.[45]

- W2W1PAR. This weight accounts for (1) school nonresponse at the base year, (2) student nonresponse in the base year and the first follow-up, (3) subsampling of parents for the first follow-up, and (4) parent nonresponse at the base year and the first follow-up. All records for sample students who participated in the base year *and* first follow-up with parents who also responded in the base year *and* first follow-up will have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of ninth-grade students. This weight can be used for analysis of population change from the base year to the first follow-up in the home-life (contextual) responses obtained from the parent questionnaires, alone or in combination with student survey and/or assessment data, administrator/counselor data from either round of HSLS:09, or teacher data from the base year.[46]

6.3 Student Weights

The cumulative HSLS:09 analysis file generated from the base year and first follow-up contains three student analytic weights for analysis involving student questionnaire responses and assessment scores.

The student analytic weight (W1STUDENT) constructed in the base year for responding study-eligible ninth-grade students is briefly reviewed in section 6.3.1. The corresponding student analytic weight constructed for the study-eligible ninth-grade students who responded to the HSLS:09 first follow-up (W2STUDENT) is detailed in section 6.3.2. The student longitudinal analytic weights (W2W1STU), created only for those students who responded to both the base year and first follow-up, are discussed in section 6.3.3. The corresponding BRR student weights are discussed in section 6.3.4.

6.3.1 Student Base-year Analytic Weights

The student base-year analytic weights were derived as a function of the school-level analytic weight and the conditional student base weight (inverse probability of selection within each school) adjusted for nonresponse and other factors. In summary, the final school-level analytic weight (W1SCHOOL) was calculated as

[45] Note that estimates generated with student data and W2W1STU in conjunction with the base-year teacher responses are no longer associated with the HSLS:09 target population of ninth-grade students and should be used with caution.

[46] Note that estimates generated with student data and W2W1PAR in conjunction with the base-year teacher responses are no longer associated with the HSLS:09 target population of ninth-grade students and should be used with caution.

$$w_{3hi} = \begin{cases} d_{hi}\, a_{1hi}\, a_{2hi}\, a_{3hi}, & \text{for eligible, responding schools} \\ d_{hi}\, a_{1hi}\, a_{3hi}, & \text{for ineligible schools} \\ 0, & \text{for eligible, nonresponding schools} \\ & \text{or hold-sample schools never released} \end{cases} \tag{6.1}$$

where d_{hi} is the base weight for school *i* in sampling stratum *h* (i.e., school *hi*), a_{1hi} is the adjustment for the random release of a portion of the complete sample of schools, a_{2hi} is the nonresponse adjustment among the study-eligible schools, and a_{3hi} is the calibration adjustment to study-eligible school counts tabulated from the 2007–08 Common Core of Data (CCD) and 2007–08 Private School Universe Survey (PSS).[47]

HSLS:09 ninth-grade students were randomly selected from four race/ethnicity sampling strata (Hispanic, Asian, Black, and Other) within each sampled school. The final student base-year analytic weight (W1STUDENT) was constructed as

$$w_{4hij} = \begin{cases} w_{3hi}\, d_{j|hi}\, a_{1hij}\, a_{2hij}\, a_{3hij}, & \text{for participating students} \\ w_{3hi}\, d_{j|hi}\, a_{3hij}, & \text{for questionnaire-incapable students} \\ 0, & \text{all other study-eligible sampled students} \end{cases} \tag{6.2}$$

where w_{3hi} is the school analytic weight defined in expression (6.1), $d_{j|hi}$ is the conditional student-level base weight (inverse probability of selection in stratum *j* within sample school *hi*), a_{1hij} is the nonresponse adjustment to account for parent refusal on behalf of the student, a_{2hij} is the nonresponse adjustment to account for direct student refusal to participate in the study, and a_{3hij} is the calibration adjustment to the study-eligible student counts tabulated from the 2007–08 CCD and 2007–08 PSS.[48]

6.3.2 Student First Follow-up Analytic Weight (W2STUDENT)

A single set of analytic weights was constructed for analyzing student data collected in the HSLS:09 first follow-up. The adjustment procedures discussed below follow those implemented in the base year. Specifically, the base weight generated in the HSLS:09 base year (section 6.3.2.1) was adjusted for two possible occurrences of nonresponse (section 6.3.2.2) and then calibrated to population totals (section 6.3.2.3).

---

[47] Weight calibration is a model-based technique used to adjust survey weights. Model covariates must be known for all respondents and may be binary, categorical, or continuous. Furthermore, population totals for all model covariates are required. Poststratification and raking are specific forms of weight calibration used in other NCES studies. Additional information on weight adjustment may be found in chapters 13 and 14 of Valliant, Dever, and Kreuter (2013).

[48] See appendix D of the HSLS:09 base-year documentation for a detailed discussion of the base-year analytic weight.

6.3.2.1 Base Weight

The student base weight developed for the HSLS:09 base year also served as the first follow-up base weight. Specifically, the HSLS:09 base weight was calculated as

$$w_{1hij} = w_{3hi} \, d_{j|hi} \,, \tag{6.3}$$

where w_{3hi} is the school analytic weight defined in expression (6.1) and $d_{j|hi}$ is the conditional student-level base weight (inverse probability of selection in stratum *j* within sample school *hi*). Note that the base weight is shown as a component of the final base-year weight given in expression (6.2).

6.3.2.2 Adjustments for Nonresponse for W2STUDENT

Table 48 shows that, among the 25,184 sampled students who remained eligible for the HSLS:09 first follow-up, 20,594 (or 81.8 percent unweighted) participated in the first follow-up.[49] Within the 4,590 nonresponding set, 1,719 (or 37.5 percent unweighted) students also did not participate in the HSLS:09 base year (double nonrespondents). The remaining 2,871 (or 62.5 percent unweighted) students only participated in the base-year study (first follow-up nonrespondents). Because different levels of information were available for the nonresponse adjustment within the two groups,[50] two nonresponse adjustments were applied to the base weight similar to those developed for the base year.

The ability of the student to participate in the study was assessed prior to first follow-up data collection, regardless of the student's status in the base year. As in the base year, students were classified as questionnaire-incapable only if they had physical limitations such as visual impairment, cognitive disabilities, or limited English proficiency. The questionnaire-incapable students were excused from taking the student survey and assessment; participation was requested from all questionnaire-capable students. All first follow-up nonresponding students thus retained the questionnaire-capable classification. These records were combined with the first follow-up responding student records for the W2STUDENT nonresponse adjustment. All questionnaire-incapable student records were excluded from this procedure, a practice that is consistent with the HSLS:09 base year.

[49] Note that the student's ability to participate in the study was determined prior to data collection. Thus, all nonresponding students were classified as questionnaire capable. Because the questionnaire-incapable students were fundamentally different from the responding and nonresponding students, they were excluded from the weight adjustment procedures.

[50] Unlike the first nonrespondent set, the second adjustment was produced with information obtained from the base-year questionnaire along with variables available on the sampling frame (see chapter 3). Limited data other than sampling information were available for the first nonrespondent set.

Table 48. Study-eligible student participation categories for HSLS:09 first follow-up: 2012

Description	Number	Percent[1]
Total number of sample students	25,206	100.0
Study-ineligible students[2]	22	0.1
Total study-eligible students	25,184	99.9
Responding student	20,594	81.8
Total nonresponding students[3]	4,590	18.2
Nonresponding students—base-year participants	2,871	62.5
Nonresponding students—base-year nonrespondents	1,719	37.5

[1] The unweighted percent for study-ineligible, questionnaire-incapable, and questionnaire-capable students is based on overall total number of sampled students. The unweighted percent for responding and nonresponding questionnaire-capable students is based on total number of questionnaire-capable students. The unweighted percent for base-year participants and nonrespondents who did not respond in the first follow-up is based on the total number of first follow-up nonresponding students. Percentages may not sum to 100 because of rounding.

[2] These students were not eligible for HSLS:09 first follow-up because they were either not in the ninth grade in the base year or died after being sampled for the base year.

[3] Questionnaire-incapable students were classified as nonrespondents for this analysis.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Both nonresponse adjustments were calculated in SUDAAN with the WTADJUST procedure using design-based logistic models.[51] Classification and regression tree (CART) analyses were used to identify important variables and combinations of those variables linked to the response patterns within the two nonresponding student sets. Additional details of each adjustment are discussed below.

Double nonrespondents. The CART analysis identified school characteristics (e.g., school type, region) and student characteristics (e.g., sex, race/ethnicity) available on the sampling frame as important variables for the first of two adjustments. The student weight adjusted for the first of two nonresponse conditions was defined as

$$w_{2hij} = \begin{cases} w_{1hij}\ a_{1hij}, & \text{for students who participated in at least the base-year study} \\ w_{1hij}, & \text{for questionnaire-incapable students} \\ 0, & \text{for students who did not respond in either round of HSLS:09} \end{cases} \tag{6.4}$$

where a_{1hij} is the first nonresponse weight adjustment that accounts for nonresponse to both HSLS:09 base year and first follow-up. Summary statistics for the first nonresponse weight adjustment are the following: minimum = 1.00, median = 1.05, and maximum = 1.34.

First follow-up nonrespondents. The positive weight in expression (6.4) was then adjusted to account for those students who participated in the base year but did not respond to the first follow-up. These include, for example, students who completed an insufficient number of

[51] See http://www.rti.org/sudaan/.

questions on the instrument to be classified as a usable case and those who were otherwise eligible but did not participate after multiple contact attempts. A second model was constructed to inflate w_{2hij} for the 20,594 eligible, responding students (table 49), resulting in the nonresponse-adjusted student weight:

$$w_{3hij} = \begin{cases} w_{2hij}\, a_{2hij}, & \text{for first follow-up participating students} \\ w_{2hij}, & \text{for questionnaire-incapable students} \\ 0, & \text{for students who did not participate in the first follow-up} \end{cases} \quad (6.5)$$

where a_{2hij} is the second student nonresponse weight adjustment calculated with SUDAAN®. Relevant model covariates included school-level geographic information, student demographic characteristics, and responses to the base-year study associated with educational aspirations and abilities. The minimum, median, and maximum values for a_{2hij} are 1.00, 1.09, and 2.84, respectively.

6.3.2.3 Weight Calibration to Produce the Final Analytic Weight W2STUDENT

A final weight adjustment was applied to the weight in expression (6.5) to maintain consistency with the student population first defined in the HSLS:09 base year. A calibration model was developed that included both school- and student-level characteristics. As with the nonresponse adjustments, the calibration adjustment was constructed and applied through the WTADJUST procedure. The final first follow-up student analytic weight (W2STUDENT) was calculated as:

$$w_{4hij} = \begin{cases} w_{3hij}\, a_{3hij}, & \text{for first follow-up participating students,} \\ & \quad \text{or for first follow-up questionnaire-incapable students} \\ 0, & \text{for nonresponding, questionnaire-capable students} \end{cases} \quad (6.6)$$

where a_{3hij} is the calibration adjustment determined through the exponential model. The minimum, median, and maximum values for a_{3hij} are 0.61, 1.19, and 4.54, respectively.

The summary statistics for the student analytic weight (W2STUDENT) in the HSLS:09 cumulative public-use file are provided below. Details of the student weight, including average calibration adjustment and sum of the final weight, are provided in table 49 by important study characteristics.

Statistic	Value
Mean	201.8
Median	138.1
Standard deviation	255.1
Minimum	2.0
Maximum	6,613.9

Table 49. Average calibration adjustments, weight sums, and unequal weighting effect, by school and student characteristics for student analytic weight (W2STUDENT): 2012

Characteristics	Number of responding students[1]	Average calibration adjustment	Final student analytic weight: Sum of the weights[2]	Final student analytic weight: Unequal weighting effect[3]
Total	20,594	1.29	4,155,676	2.60
School type				
Public	16,845	1.30	3,858,126	2.41
Private	3,749	1.24	297,551	2.02
Region				
Northeast	3,208	1.20	723,085	3.81
Midwest	5,501	1.33	919,570	1.90
South	8,432	1.27	1,562,209	2.10
West	3,453	1.36	950,813	2.50
Locale				
City	5,852	1.33	1,329,020	3.81
Suburban	7,378	1.29	1,384,714	1.88
Town	2,447	1.20	484,272	2.09
Rural	4,917	1.28	957,671	1.82
Student sex				
Male	10,384	1.28	2,088,375	2.39
Female	10,210	1.30	2,067,302	2.81
Student race/ethnicity[4]				
Hispanic	3,271	1.30	928,628	3.12
Asian	1,675	1.27	147,067	3.73
Black	2,121	1.32	569,991	3.15
Other	13,527	1.28	2,509,990	1.77

[1] The questionnaire-incapable students have been excluded from the analysis presented in this table. The sum of W2STUDENT on the HSLS:09 restricted-use file is 4,197,724.

[2] The student counts in table 10 of chapter 3 within the base-year documentation (Ingels et al. 2011) were used as the control totals. Weight sums differ from the population counts because of the suppression of the questionnaire-incapable students from the public-use file. Values may not sum to overall total because of rounding.

[3] The unequal weighting effect is also referred to as the design effect of the weights and is calculated as one plus the square of the coefficient of variation ($1 + CV^2$).

[4] Variable = X2RACE where "Other" includes White and other race/ethnicities.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

6.3.3 Student Longitudinal Analytic Weight (W2W1STU)

A set of analytic weights was constructed for analyzing change in student responses collected in the HSLS:09 base-year and first follow-up instruments. The analytic weight, W2STUDENT given in expression (6.6), was adjusted to address absence of student data for those who did not participate in the base year but did participate in the HSLS:09 first follow-up (table 50). In other words, a positive longitudinal weight was constructed only for students who participated in the base year *and* first follow-up. Specifically, the student first follow-up analytic weight (section 6.3.2) were adjusted for nonresponse associated with the base year (section 6.3.3.1) and then calibrated to the student population counts (section 6.3.3.2).

Table 50. Study-eligible student participation categories for HSLS:09 base year, by first follow-up: 2012

Base-year student participation category	First follow-up participation category[1] Total		Study ineligible		Responding students		Nonresponding students	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	25,206	100.0	22	0.1	20,594	81.7	4,590	18.2
Responding student	21,444	85.1	15	0.1	18,623	73.9	2,806	11.1
Nonresponding students	3,762	14.9	7	0.0	1,971	7.8	1,784	7.1

[1] The unweighted percentages in the Total column are based on the overall total number of base-year-eligible, sampled students (n = 25,206). The unweighted percent within the first follow-up participation categories is based on the total column within each row. Percentages may not sum to 100 because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

6.3.3.1 Adjustments for Nonresponse for W2W1STU

A CART analysis was used to construct the weighting classes for the nonresponse adjustment. The variables included in this analysis were the same as those used in the second nonresponse adjustment applied to construct W2STUDENT (section 6.3.2.2). The longitudinal student weight adjusted for nonresponse in the base year was constructed as

$$w_{5hij} = \begin{cases} w_{4hij}\; a_{4hij}, & \text{for students who responded to the base year and first follow-up,} \\ & \quad \text{or for questionnaire incapable base-year students who responded} \\ & \quad \text{in the first follow-up} \\ 0, & \text{for students who refused to participate in either base year} \\ & \quad \text{or first follow-up} \end{cases} \tag{6.7}$$

where a_{4hij} is the nonresponse adjustment determined through the WTADJUST logistic model and w_{4hij} is W2STUDENT, the student first follow-up analytic weight described in section 6.3.2. The minimum, median, and maximum values for a_{4hij} are 1.00, 1.01, and 6.51, respectively.

6.3.3.2 Weight Calibration to Produce the Final Analytic Weight W2W1STU

A calibration factor was applied to the nonresponse-adjusted weight given in expression (6.7) to produce the final set of student longitudinal analytic weights. As with the other weights, the calibration adjustment was formed with an exponential model that included characteristics for the sampled schools and students. The student longitudinal analytic weight (W2W1STU) was calculated as:

$$w_{6hij} = \begin{cases} w_{5hij}\, a_{5hij}, & \text{for students who responded to the base year } \textit{and} \text{ first follow-up,} \\ & \text{or for questionnaire incapable base-year students who responded} \\ & \quad \text{in the first follow-up,} \\ & \text{or for students designated as questionnaire incapable in both the} \\ & \quad \text{base year } \textit{and} \text{ first follow-up} \\ 0, & \text{for students who refused to participate in either base year} \\ & \quad \text{or first follow-up} \end{cases} \tag{6.8}$$

where a_{5hij} is the calibration adjustment determined through the exponential model and w_{5hij} is W2STUDENT, the student first follow-up analytic weight, further adjusted for nonresponse in the base year (see expression 6.7). The minimum, median, and maximum values for a_{5hij} are 0.62, 1.19, and 4.58, respectively.

The summary statistics for the student longitudinal analytic weight (W2W1STU) in the HSLS:09 cumulative public-use file are provided below. Details of the student longitudinal weight, including average calibration adjustment and sum of the final weight, are provided in table 51 by important school-level characteristics.

Statistic	Value
Mean	220.3
Median	146.3
Standard deviation	284.9
Minimum	1.9
Maximum	7,648.3

Table 51. Average calibration adjustments, weight sums, and unequal weighting effect, by school and student characteristics for student longitudinal weight (W2W1STU): 2012

Characteristics	Number of responding students[1]	Average calibration adjustment	Final student analytic weights: Sum of the weights[2]	Final student analytic weights: Unequal weighting effect[3]
Total	18,623	1.29	4,101,850	2.67
School type				
Public	15,164	1.29	3,805,209	2.47
Private	3,459	1.25	296,641	2.10
Region				
Northeast	2,894	1.20	710,116	3.88
Midwest	4,997	1.31	912,909	2.02
South	7,592	1.27	1,541,644	2.19
West	3,140	1.36	937,181	2.56
Locale				
City	5,259	1.33	1,309,428	3.85
Suburban	6,577	1.28	1,366,338	2.03
Town	2,247	1.21	479,419	2.17
Rural	4,540	1.28	946,665	1.83
Student sex				
Male	9,349	1.28	2,062,186	2.46
Female	9,274	1.29	2,039,663	2.89
Student race/ethnicity[4]				
Hispanic	2,958	1.29	913,891	3.16
Asian	1,467	1.27	140,374	3.94
Black	1,898	1.31	563,487	3.31
Other	12,300	1.28	2,484,098	1.83

[1] The questionnaire-incapable students have been excluded from the analysis presented in this table. The sum of W2W1STU on the HSLS:09 restricted-use file is 4,197,724.
[2] The student counts in table 5 of chapter 3 within the base-year documentation (Ingels et al. 2011) were used as the control totals Weight sums differ from the population counts because of the suppression of the questionnaire-incapable students from the public-use file. Values may not sum to overall total because of rounding.
[3] The unequal weighting effect is also referred to as the design effect of the weights and is calculated as one plus the square of the coefficient of variation ($1 + CV^2$).
[4] Variable = X2RACE where “Other” includes White and other race/ethnicities.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

6.3.4 Balanced Repeated Replication Weights

A set of 200 BRR weights was created for school-level analyses in the base year and is available on the cumulative analytic file. In addition, three sets of 200 student-level BRR weights were developed: (1) base-year student BRR weights (W1STUDENT001–200); (2) first follow-up student BRR weights (W2STUDENT001–200); and (3) base-year to first follow-up student longitudinal weights (W2W1STU001–200). Procedures for constructing the weights mirrored those used to construct the corresponding analytic weight. Namely, BRR base weights were constructed and subjected to nonresponse and calibration adjustments developed for each replicate. The procedures for creating the BRR weights are discussed in section 6.4.4.

6.4 Student Contextual Weights

Contextual information was requested for all sampled students in the base year (administrators, counselors, parents, science teachers and mathematics teachers) and for some sampled students in the first follow-up (administrators, counselors, and parents). See chapter 3 for additional details on the sample design for the contextual data. However, not all persons identified to provide this information agreed to participate in HSLS:09. When the response rates were judged to be low, additional weights were created for analyzing HSLS:09 data that also include responses from the contextual instruments. The base-year contextual weights are summarized in section 6.4.1, followed by first follow-up home-life contextual weights in section 6.4.2 and the longitudinal home-life contextual weights in section 6.4.3.

6.4.1 Base-year Student Contextual Weights

Three contextual weights were calculated for the HSLS:09 base year and are included on the cumulative file.[52] The contextual weights are:

- Science Course Enrollee Contextual Weight (W1SCITCH),
- Mathematics Course Enrollee Contextual Weight (W1MATHTCH), and
- Student Home-life Contextual Weights (W1PARENT).

These weights were generated to account for differential and non-negligible nonresponse patterns within the groups of adults providing contextual information. However, as discussed in the HSLS:09 base-year documentation, the calibration adjustments—applied to the student analytic weights (W1STUDENT) to produce the contextual weights—only used student characteristics because such information was not available for all teachers and parents of the sampled students. Note that contextual data were also collected from school administrators and counselors but, because of the high response rates, analyses including this information are implemented with the student weights (see section 6.2).

[52] Base-year contextual weights were not generated for analysis of administrator and counselor data because of the high response rates.

This same approach was applied independently to each base-year student BRR weight (W1STUDENT001–200) to calculate the corresponding BRR contextual weights (i.e., W1SCITCH001–200, W1MATHTCH001–200, and W1PARENT001–200, respectively).

6.4.2 First Follow-up Student Contextual Analytic Weight (W2PARENT)

Details on family and home-life conditions and the importance of education were again requested from the parent/guardian of the sampled students. In contrast to the base year, where contextual information from parents and/or guardians was attempted for the full sample of students, contextual information in the first follow-up was collected for a 47.4 percent random subsample of students (table 52).

Table 52. Parent participation categories for HSLS:09 first follow-up study-eligible students: 2012

Description	Number	Unweighted percent[1]	Weighted percent[2]
Total number of sample students	25,206	100.0	100.0
Study-ineligible students[3]	22	0.1	0.1
Total eligible students not in parent subsample	13,232	52.5	51.2
Total eligible students in parent subsample	11,952	47.4	48.7
Responding parents	8,651	72.4	72.5
Nonresponding parents	3,301	27.6	27.5

[1] The unweighted percent for study ineligibles and inclusion (or not) in the parent subsample is based on overall total number of sampled students (n = 25,206). The unweighted percent by parent response status is based on total number of study-eligible students in the parent subsample. Percentages may not sum to 100 because of rounding.
[2] Weighted percentages for responding and nonresponding parent rows were calculated using the student base weight that was adjusted for the random parent subsampling. Other weighted percentages were calculated with the student base weight directly.
[3] These students were not eligible for HSLS:09 first follow-up because they were either not in the ninth grade in the base year or died after being sampled for the base year.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

The base weight for W2PARENT was constructed to reflect the subsampling of cases:

$$w_{7hij} = \begin{cases} w_{1hij}\, a_{6hij}, & \text{for students randomly selected for the first follow-up parent subsample} \\ 0, & \text{for students not selected for the subsample, or study-ineligible students identified prior to sampling} \end{cases} \qquad (6.9)$$

where a_{6hij} is the inverse of the subsampling rate described in section 3.3.4, and w_{1hij} is the first follow-up student base weight defined in expression (6.3). Summary statistics for the W2PARENT weight adjustment are the following: minimum = 1.00, median = 1.17, and maximum = 5.

Among the 11,952 subsampled students (table 52), home-life contextual (parent interview) responses were obtained from 8,651 parents (72.4 percent unweighted; 72.5 percent weighted). A low to moderate response rate is one measure that might suggest the need for a separate analytic weight that has been adjusted to lower potential nonresponse bias. However, because parent information was not available for all sampled students, nonresponse bias associated with parental response patterns could not be evaluated and adjusted for directly. Instead, the contextual weights were formed with an adjustment using school and student information in a calibration (exponential) model. In other words, the first follow-up student home-life contextual weights (W2PARENT) were developed by calibrating the student base weights adjusted for subsampling (i.e., expression 6.10), to the population counts used for the other analytic weights:

$$w_{8hij} = \begin{cases} w_{7hij}\, a_{7hij}, & \text{for student subsample cases with a responding parent} \\ 0, & \text{for student subsample cases without a responding parent,} \\ & \text{or students not selected for the parent subsample,} \\ & \text{or study-ineligible students identified prior to subsampling} \end{cases} \quad (6.10)$$

where a_{7hij} is the calibration adjustment determined through an exponential model including school and student characteristics, and w_{7hij} is student base weight adjusted for subsampling given in expression (6.9). The minimum, median, and maximum calibration adjustments were 0.85, 2.28, and 8.74, respectively.

The summary statistics for the student home-life contextual weight (W2PARENT) in the cumulative HSLS:09 public-use file are given below. Details of the parent longitudinal weights, including average calibration adjustment and sum of the final weight, are provided in table 53 by important school-level characteristics.

Statistic	Value
Mean	478.2
Median	329.5
Standard deviation	591.3
Minimum	3.0
Maximum	17,773.6

Table 53. Average calibration adjustments, weight sums, and unequal weighting effect, by school and student characteristics for student home-life contextual weight (W2PARENT): 2012

Characteristics	Number of responding parents[1]	Average calibration adjustment	Final analytic weight: Sum of the weights[2]	Final analytic weight: Unequal weighting effect[3]
Total	8,621	1.81	4,122,897	2.53
School type				
Public	7,091	1.84	3,825,142	2.36
Private	1,530	1.69	297,755	2.04
Region				
Northeast	1,350	1.71	715,209	3.80
Midwest	2,228	1.89	916,451	2.03
South	3,490	1.76	1,542,219	2.09
West	1,553	1.91	949,018	2.34
Locale				
City	2,520	1.87	1,309,318	3.71
Suburban	3,140	1.86	1,382,699	1.79
Town	982	1.60	482,859	2.09
Rural	1,979	1.77	948,021	1.93
Student sex				
Male	4,364	1.80	2,068,514	2.44
Female	4,257	1.83	2,054,384	2.62
Student race/ethnicity[4]				
Hispanic	1,401	1.91	915,675	3.01
Asian	664	1.97	139,168	3.30
Black	948	1.91	562,022	3.27
Other	5,608	1.75	2,506,033	1.79

[1] The questionnaire-incapable students have been excluded from the analysis presented in this table. The sum of W2PARENT on the HSLS:09 restricted-use file is 4,197,724.

[2] The student counts in table 10 of chapter 3 within the base-year documentation (Ingels et al. 2011) were used as the control totals. Weight sums differ from the population counts because of the suppression of the questionnaire-incapable students from the public-use file. Values may not sum to overall total because of rounding.

[3] The unequal weighting effect is also referred to as the design effect of the weights and is calculated as one plus the square of the coefficient of variation ($1 + CV^2$).

[4] Variable = X2RACE where “Other” includes White and other race/ethnicities.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Note that unlike the base year, the final home-life contextual weights were constructed for all student cases with a responding parent, regardless of the student’s response status in the first follow-up.[53] This change produced 323 student records with a positive W2PARENT weight but a W2STUDENT equal to zero (table 54).

[53] Home-life contextual weights in the base year were constructed for sampled cases having both a responding student *and* a responding parent.

Table 54. Student and parent participation categories for HSLS:09 first follow-up study-eligible student cases: 2012

Description	Number	Percent[1]
Total number of students in parent subsample[2]	11,952	100.0
Responding parents	8,651	72.4
Responding students	8,328	96.3
Nonresponding students	323	3.7
Nonresponding parents	3,301	27.6
Responding students	1,552	47.0
Nonresponding students	1,749	53.0

[1] The unweighted percent by parent response status is based on overall total number of subsampled students. The unweighted percent by student response status is based on the total number of subsampled students within parent response.

[2] Nine study-ineligible students are excluded from the subsample counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

Other First Follow-up Student Contextual Weights. First follow-up contextual information was collected from three sources for the sampled students. As detailed in chapter 4, questionnaires were sent to all school administrators (base-year and transfer schools) and a school counselor at the base-year schools, along with the parents of students randomly selected for the parent subsample discussed above. Administrators for base-year and transfer schools and counselors at the base-year schools only were asked to comment on the climate of the school. As in the base year, the contextual responses were obtained for a high proportion of students in the sample (see the staff coverage rates in table 16, chapter 4). Thus, separate contextual weights were not constructed for the administrator and counselor data. Instead, as discussed in section 6.2, other weights are used for analyses that include these responses.

6.4.3 Longitudinal Student Home-life Contextual Weights (W2W1PAR)

As with change in student responses (section 6.3.3), analytic weights were also constructed for analyzing change in home-life characteristics and opinions from the time of the HSLS:09 base year to first follow-up in tandem with changes to the student responses. As shown in table 55, 6,371 records (or 53.3 percent of the 11,952 subsample cases) contained student and parent responses in the base year and first follow-up. Therefore, a longitudinal student home-life contextual weight (W2W1PAR) was created for these cases.

Table 55. Parent participation categories for HSLS:09 base year, by first follow-up: 2012

Description	Number	Percent[1]
Total students in parent subsample[2]	11,952	100.0
Subsample students who did not respond in base year or first follow-up[3]	2,940	24.6
Subsample students who responded in base year *and* first follow-up	9,012	75.4
First follow-up nonresponding parents	1,339	14.9
First follow-up responding parents	7,673	85.1
Base-year responding parents	6,371	83.0
Base-year nonresponding parents	1,302	17.0

[1] The unweighted percent by parent response status is based on overall total number of subsampled students. The unweighted percent by student response status is based on the total number of subsampled students within parent response.

[2] Nine study-ineligible students are excluded from the subsample counts.

[3] The questionnaire-incapable students are included in the set who did not respond to the base year and first follow-up.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

The longitudinal analytic weight for the contextual data, W2W1PAR, was created by modifying the first follow-up contextual base weight created for W2PARENT, expression (6.9), to account for the loss of a total of 5,581 cases: (1) 2,940 cases without student response in both rounds of HSLS:09, (2) 1,339 cases without parent response in the first follow-up, and (3) 1,302 cases without parent response in the base year. Specifically, W2W1PAR was constructed as

$$w_{9hij} = \begin{cases} w_{7hij}\, a_{8hij}, & \text{for student subsample cases with parent responses in} \\ & \quad \text{the HSLS:09 base year } \textit{and} \text{ first follow-up} \\ 0, & \text{for student subsample cases without parent responses in} \\ & \quad \text{the HSLS:09 base year } \textit{and} \text{ first follow-up,} \\ & \quad \text{or students not selected for the parent subsample,} \\ & \quad \text{or study-ineligible students identified prior to subsampling} \end{cases} \qquad (6.11)$$

where a_{8hij} is the calibration (exponential model) adjustment using school and student characteristics to account for cases without student and parental responses in the base year and first follow-up, and w_{7hij} is student base weight adjusted for subsampling given in expression (6.8). The minimum, median, and maximum calibration adjustments were 0.85, 2.28, and 8.74, respectively.

The summary statistics for the student home-life longitudinal weight (W2W1PAR) in the cumulative HSLS:09 public-use file are provided below. Details of the student home-life longitudinal weights, including average calibration adjustment and sum of the final weight, are provided in table 56 by important school and student characteristics.

Statistic	Value
Mean	650.9
Median	438.9
Standard deviation	861.2
Minimum	6.5
Maximum	25,718.0

Table 56. Average calibration adjustments, weight sums, and unequal weighting effect, by school and student characteristics for student home-life longitudinal weight (W2W1PAR): 2012

Characteristics	Number of responding parents[1]	Average calibration adjustment	Final student analytic weights	
			Sum of the weights[2]	Unequal weighting effect[3]
Total	6,371	2.47	4,146,720	2.75
School type				
Public	5,133	2.56	3,849,099	2.52
Private	1,238	2.10	297,621	2.09
Region				
Northeast	1,005	2.32	722,349	4.67
Midwest	1,672	2.58	919,703	2.00
South	2,564	2.43	1,558,826	2.24
West	1,130	2.54	945,843	2.36
Locale				
City	1,870	2.52	1,318,970	4.31
Suburban	2,322	2.55	1,381,701	1.83
Town	714	2.22	487,953	2.16
Rural	1,465	2.41	958,095	1.89
Student sex				
Male	3,197	2.48	2,093,266	2.66
Female	3,174	2.46	2,053,454	2.84
Student race/ethnicity[4]				
Hispanic	1,013	2.70	923,333	3.20
Asian	483	2.71	141,577	3.58
Black	651	2.77	548,551	3.88
Other	4,224	2.34	2,533,259	1.83

[1] The questionnaire-incapable students have been excluded from the analysis presented in this table. The sum of W2W1PAR on the HSLS:09 restricted-use file is 4,197,724.
[2] The student counts in table 5 of chapter 3 within the base-year documentation (Ingels et al. 2011) were used as the control totals Weight sums differ from the population counts because of the suppression of the questionnaire-incapable students from the public-use file. Values may not sum to overall total because of rounding.
[3] The unequal weighting effect is also referred to as the design effect of the weights and is calculated as one plus the square of the coefficient of variation ($1 + CV^2$).
[4] Variable = X2RACE where the “Other” includes White and other race/ethnicities.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up.

6.4.4 Balanced Repeated Replication Weights

Similar to the student-level BRR weights, procedures used to develop the analytic weights were replicated independently for each weight within five sets of 200 BRR weights. These included one set of base-year BRR weights for analyses including parent responses; one set each for use with data obtained from the science and mathematics teacher questionnaires; one set of first follow-up home-life contextual BRR weights; and one set of home-life contextual longitudinal BRR weights. The procedures for creating the weights are discussed in the next section. The variable names for the weights included on the cumulative data file are given in table 57.

Table 57. HSLS:09 analytic and BRR weights, by type of HSLS:09 analysis

Type of HSLS:09 analysis		Variable names for HSLS:09 analytic weights	Variable names for HSLS:09 BRR weights
Base year	School-level	W1SCHOOL	W1SCHOOL001–W1SCHOOL200
	Student-level	W1STUDENT	W1STUDENT001–W1STUDENT200
	Contextual analyses	W1PARENT	W1PARENT001–W1PARENT200
		W1SCITCH	W1SCITCH001–W1SCITCH200
		W1MATHTCH	W1MATHTCH001–W1MATHTCH200
First follow-up	Student-level	W2STUDENT	W2STUDENT001–W2STUDENT200
	Contextual analyses	W2PARENT	W2PARENT001–W2PARENT200
Base year and first follow-up	Student-level	W2W1STU	W2W1STU001–W2W1STU200
	Contextual analyses	W2W1PAR	W2W1PAR001–W2W1PAR200

NOTE: BRR = balanced repeated replication.
SOURCE: U.S. Department of Education, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Data File.

6.5 Variance Estimation

Analysis of HSLS:09 data requires statistical software that can calculate either (a) balanced repeated replication (BRR) variance estimates using the BRR weights and the associated analytic weight, or (b) linearization variance estimates through a Taylor series approximation using only the analytic weight.[54] Some standard software packages, however, do not calculate estimates that account for the random sampling of students clustered within schools. This incorrect design assumption can lead to estimated variances and confidence intervals that are too small and this may lead to incorrect results from hypothesis tests. Therefore, researchers are advised to use appropriate software such as SUDAAN and Stata and are provided with example code in the next section.

[54] National Center for Education Statistics (NCES) standards recommend the use of replicate variance estimation over linearization methods. The sample design variables, strata and primary sampling units (PSU), were suppressed from the public-use file as one measure of disclosure avoidance (see section 7.5 for the disclosure risk analysis and protection).

The importance of correct variance estimation is further emphasized in section 6.5.1 through a discussion of the BRR and linearization methodologies. The variance inflation associated with the clustered HSLS:09 sample design and associated weights in comparison to an unclustered design, quantified in the design effect, is discussed in section 6.5.2.

6.5.1 Standard Errors

The two methods of variance estimation available for HSLS:09 are BRR and Taylor series linearization. BRR variance estimation is available with either the HSLS:09 restricted-use or public-use files. This method does not need the analytic stratum and PSU identifiers but does require a large set of replicate weights along with the associated analytic weight. As discussed in the HSLS:09 base year documentation (Ingels et al. 2011), the replicate weights account for several random processes including sampling and weighting and produce estimates that are in general slightly larger than the corresponding estimates calculated with linearization (Wolter 2007).

To create the school BRR weights, the original analytic strata were collapsed into 199 BRR strata with representation across the characteristics used in sampling (i.e., school type, region, and locale) and two BRR PSUs were formed. The BRR strata were randomly assigned to rows of a 200 × 200 Hadamard matrix containing a sequence of +1 and -1 values that were used to form BRR base weights. The base weights were then adjusted using procedures similar to those implemented for the analytic weights.

The general formula for calculating a BRR variance estimate, used in software packages designed for survey estimation, is:

$$var\left(\hat{\theta}\right)=\frac{1}{200}\sum_{a=1}^{200}\left(\hat{\theta}_{(a)}-\hat{\theta}\right)^{2}$$

where 200 is the number of HSLS:09 BRR weights, $\hat{\theta}$ is the estimated value for a statistic of interest (e.g., mean) calculated with a particular analytic weight, and $\hat{\theta}_{(a)}$ is the corresponding value calculated with the *a*th BRR (replicate) weight (a = 1, . . . ,200).

Taylor Series linearization variance estimation requires software that constructs a first-order Taylor-series approximation of the statistic being analyzed (e.g., mean), and data sources containing the relevant analytic weight (table 48) and the analytic stratum and primary sampling unit (PSU) identifiers (see, e.g., Binder [1983]; Woodruff [1971]). The PSU variable STRAT_ID is a unique value randomly generated for each sampled school. The 450 analytic strata were constructed in the base year by combining two to three schools into one stratum in such a way as to maximize retention of the original two-stage sample design and also precision of the estimates through the degrees of freedom (Chromy 1981). To lower disclosure risk, linearization variance estimation is only permitted through the HSLS:09 restricted-use file which, unlike the public-use file, contains the stratum and PSU variables.

Software currently available for survey data analysis includes SUDAAN®, SAS® survey procedures,[55] WesVar®,[56] Stata®,[57] R®,[58] and SPSS®.[59] Example SUDAAN code for producing estimated means and standard errors using the linearization and BRR methods are shown in figures 8 and 9, respectively. The corresponding Stata code is provided in figures 10 and 11.

Figure 8. Example SAS-SUDAAN code to calculate an estimated mean and linearization standard error for a base-year student-level analysis

```
PROC SORT DATA=<filename>;                         *File sorted by nest variables;
  BY STRAT_ID PSU;
RUN;

PROC DESCRIPT DATA=<filename> DESIGN=WR;
  NEST STRAT_ID PSU / MISSUNIT;                            *Analysis stratum/PSU;
  SUBPOPN (<domain variable = level>);                 *Subset to reporting domain;
  WEIGHT W1STUDENT;                                        *Main analytic weight;
  VAR <analysis variable>;                                    *Analysis variable;
  PRINT MEAN SEMEAN / STYLE=NCHS;                         *Mean and standard error;
RUN;
```

Figure 9. Example SUDAAN code to calculate an estimated mean and replicate (BRR) standard error for a first follow-up analysis of the home-life variables

```
PROC DESCRIPT DATA=<filename> DESIGN=BRR;
  WEIGHT W2PARENT;                                         *Main analytic weight;
  REPWGT W2W1PAR001- W2W1PAR200;                          *BRR replicate weights;
  SUBPOPN (<domain variable = level>);                 *Subset to reporting domain;
  VAR <analysis variable>;                                    *Analysis variable;
  PRINT MEAN SEMEAN / STYLE=NCHS;                         *Mean and standard error;
RUN;
```

NOTE: BRR = balanced repeated replication.

[55] See the most recent SAS/STAT User's Guide, located at http://support.sas.com/documentation/.
[56] See http://www.westat.com/westat/statistical_software/WesVar/index.cfm.
[57] See http://www.stata.com/.
[58] See http://www.r-project.org/.
[59] See http://www-01.ibm.com/software/analytics/spss/.

Figure 10. Example STATA code to an estimated mean and linearization standard error for a first follow-up student-level analysis

```
SVYSET PSU [PWEIGHT=W2STUDENT], STRATA (STRAT_ID) VCE(LINEAR),
singleunit(centered)
SVY, SUBP (<domain variable >) : MEAN < analysis variable >
```

Figure 11. Example STATA code to produce mean and replicate (BRR) standard error for a base-year student analysis with mathematics teacher data

```
SVYSET [PWEIGHT=W1MATHTCH], BRRWEIGHT(W1STUDENT001-W1STUDENT200) VCE(BRR)
SVY, SUBP (<domain variable >) : MEAN < analysis variable >
```

NOTE: BRR = balanced repeated replication.

Standard errors for a select number of variables are provided in appendix F along with their design effects as discussed in the next section.

6.5.2 Design Effects

Design effects (*deff*) measure the relative efficiency of a sample design using particular items collected in the survey. These values are calculated as the ratio of two estimated variances,

$$deff = \frac{\hat{V}_d\left(\hat{\theta}\right)}{\hat{V}_s\left(\hat{\theta}\right)}, \qquad (6.12)$$

for an estimated HSLS:09 characteristic $\hat{\theta}$. The numerator value, $\hat{V}_d\left(\hat{\theta}\right)$, is the estimated variance that properly accounts for the complex sample design and the variability associated with the analytic weights. The denominator value, $\hat{V}_s\left(\hat{\theta}\right)$, is the estimated variance from a simple random sample (*srs*) design of the same size.

In addition to *deff*, the root design effect or *deft* was also calculated. Like *deff*, this statistic also provides a measure of relative efficiency of a sample design but in terms of the standard errors:

$$deft = \sqrt{\frac{\hat{V}_d\left(\hat{\theta}\right)}{\hat{V}_s\left(\hat{\theta}\right)}}, \qquad (6.13)$$

where the components are the same as defined for expression (6.12).

The HSLS:09 first follow-up *deff*/*deft* analysis included 63 variables. As with the estimated standard errors, the *deff* and *deft* estimates were produced using final analytic weights

and data that were edited, imputed (if applicable), and treated to limit disclosure risk. The *deff* were calculated using a model-based formulation, $deff_4$ in SUDAAN. The estimates subject to this analysis included 37 student-questionnaire variables and the mathematics achievement score (*theta*), and 25 parent-questionnaire items. As in the base year, the items were chosen using three criteria: (1) variables common to the HSLS:09 base-year design effect analysis; (2) variables identified for the first follow-up First Look report; and (3) variables included in several other NCES studies such as ELS:2002 and the National Education Longitudinal Study of 1988 (NELS:88). The *deff* and *deft* estimates for the 63 study items within a set of important characteristics are provided in appendix F. The average *deff* and *deft* across the 63 items is presented in table 58.

Table 58. Average design effects (*deff*) and root design effects (*deft*) for student and parent data: 2012

Characteristic[1]	Student respondents	Final student weights: Average *deff*[2]	Final student weights: Average *deft*[3]	Parent respondents	Final home-life weights: Average *deff*[2]	Final home-life weights: Average *deft*[3]
Total	20,594	4.4	2.1	8,651	3.6	1.9
School type						
Public	16,845	4.0	2.0	7,120	3.3	1.8
Private	3,749	5.8	2.3	1,531	3.4	1.8
Region						
Northeast	3,208	5.2	2.2	1,355	4.7	2.1
Midwest	5,501	3.7	1.9	2,235	3.2	1.8
South	8,432	4.0	2.0	3,503	3.2	1.8
West	3,453	4.3	2.0	1,558	3.8	1.9
Locale						
City	5,852	5.9	2.4	2,530	4.8	2.2
Suburban	7,378	3.3	1.8	3,147	3.3	1.8
Town	2,447	4.0	2.0	985	3.4	1.8
Rural	4,917	3.8	1.9	1,989	3.6	1.9
Student sex						
Male	10,385	3.7	1.9	4,384	3.5	1.9
Female	10,209	3.7	1.9	4,267	3.1	1.8
Student race/ethnicity[4]						
Hispanic	3,271	4.0	2.0	1,409	3.5	1.9
Asian	1,675	4.9	2.2	667	3.7	1.9
Black	2,121	3.7	1.9	955	3.5	1.9
White	11,532	2.7	1.6	4,745	2.3	1.5
More than one race	1,995	3.1	1.7	875	2.9	1.7
Socioeconomic status[5]						
Low SES	4,051	3.3	1.8	1,957	3.4	1.8
Middle SES	12,389	3.7	1.9	4,639	2.7	1.6
High SES	4,154	2.7	1.6	2,055	2.3	1.5

[1] The characteristics presented here reflect the information obtained during HSLS:09 base year and do not contain updated information presented on the cumulative first follow-up data file to enable comparison with the base-year documentation.

[2] The formula for the design effect (*deff*) is provided in expression (6.12).

[3] The formula for the root design effect (*deft*) is provided in expression (6.13).

[4] Race/ethnicity as defined in the student questionnaire.

[5] Categories for socioeconomic status (SES) were defined using the SES quintile variable (X2SESQ5) where X2SESQ5 = 1 (20th percentile) represents low SES, and X2SESQ5 = 5 (80th percentile) represents high SES. All others were classified as middle SES.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Data File.

6.6 Unit Nonresponse Bias Analysis

NCES standards require unit nonresponse bias analyses when either overall or domain-specific weighted response rates fall below 85 percent. These analyses identify any statistically detectable differences between estimates calculated for study respondents and for the nonrespondents.

Bias reduction attributable to base weight adjustments for nonresponse is described below, beginning with a description of the statistical test for unit nonresponse bias (section 6.6.1). These include the student first follow-up analytic weights W2STUDENT (section 6.6.2.1) and student longitudinal weights W2W1STU (section 6.6.2.2). Analyses of the home-life contextual weights W2PARENT and the associated longitudinal weights W2W1PAR are summarized in sections 6.6.3.1 and 6.6.3.2, respectively. Results from the bias analyses are concluded in section 6.6.3.3 with a discussion of the student contextual responsive design.

6.6.1 Test of Significant Nonresponse Bias

Nonresponse bias is the difference between the estimated parameter calculated from the respondent data and the true value. For a population mean, for example, the nonresponse bias is calculated as

$$Bias\left(\overline{y}_R\right) = \overline{y}_R - \mu \qquad (6.14)$$

where $\overline{y}_R$ is the mean (or proportion) estimated from the survey responses and μ is the corresponding (true) population value. Because the truth is unknown, the population value and the bias must be estimated using data from respondents and nonrespondents:

$$\hat{\mu} = \left(1 - \hat{\eta}\right)\overline{y}_R + \hat{\eta}\overline{y}_{NR} \qquad (6.15)$$

where $\hat{\eta}$ is the weighted unit nonresponse rate.[60] Substituting expression (6.15) into expression (6.14) provides the formula for the estimated bias

$$\hat{B}ias\left(\overline{y}_R\right) = \hat{\eta}\left(\overline{y}_R - \overline{y}_{NR}\right) \qquad (6.16)$$

Initial bias estimates were calculated with the DESCRIPT procedure in SUDAAN and the (adjusted) base weights were used to generate the nonresponse rate. With the estimated standard error of the bias that accounted for the association between $\overline{y}_R$ and $\overline{y}_{NR}$, a *t* test was formed to determine whether the bias was significantly greater than zero at a 0.05 level of significance. The same test was recomputed using nonresponse-adjusted weights to determine whether the weight adjustment appropriately reduced the bias to insignificant levels. Table 59 contains a summary of the analysis for the four analytic weights—see table F-1 in appendix F for the detailed analysis tables.

[60] The weighted unit nonresponse rate was calculated using the design weights adjusted for school release and the student design weights for each type of nonresponse bias analysis.

Table 59. Summary statistics for unit nonresponse bias analyses before and after a weight adjustment for nonresponse, by HSLS:09 analytic weight: 2012

Analytic weight	Significant bias tests at 0.05 level[1]		Median absolute relative bias[2]		
	Percent before weight adjustment	Percent after weight adjustment	Percent before weight adjustment	Percent after weight adjustment	Percent relative change[3]
Student					
First follow-up	31.8	0	1.2	0.3	-72.3
Longitudinal	33.3	0	1.6	0.3	-79.0
Student home-life contextual					
First follow-up	25.8	0	2.3	0.2	-89.7
Longitudinal	37.9	3.0	5.0	0.5	-90.4

[1] Before and after are in reference to the nonresponse weight adjustment. A total of 66 statistical tests were performed; the number 66 was used as the basis for the reported percentages.
[2] The percent relative bias is calculated as 100 multiplied by the estimated bias divided by the estimated value. The absolute relative bias is the absolute value of the (percent) relative bias.
[3] The percent relative change is calculated as 100 multiplied by the median value after adjustment minus the median value before adjustment divided by the median value before adjustment.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Data File.

6.6.2 Nonresponse Bias Analyses of Student Data

6.6.2.1 First Follow-up Student-level Nonresponse Bias Analysis

In keeping with the NCES statistical standards, nonresponse bias analyses were performed for first follow-up student responses using the student analytic weight W2STUDENT (section 6.6.2.1) because the overall weighted response rate was 81.8 percent (table 47). Students who completed a substantial portion of the questionnaire were classified as a respondent, regardless of their level of participation in the mathematics assessment.

As in the base year, some information (e.g., race/ethnicity, sex) was available for nonresponding students through the study enrollment lists. Base-year school characteristics were available for all sampled students. Note that first follow-up school characteristics were not available or not applicable in the case of early graduates for some nonresponding students. In total, 17 variables were used for the student nonresponse bias analysis. Approximately 31.8 percent of the 66 statistical tests identified bias significantly greater than zero at the 0.05 significance level (table 59) prior to adjusting the weights for nonresponse. After adjustment, no levels of bias were detectable at the 0.05 level of significance and the median absolute relative bias was reduced by 72.3 percent. The results are presented in table F-1 in appendix F.

6.6.2.2 Longitudinal Student-level Nonresponse Bias Analysis

Nonresponse bias was also evaluated in student items available for a longitudinal analysis. As shown in table 47, the overall weighted response rate for the first follow-up was 81.8 percent. However, the overall weighted response rate for students with responses in the first follow-up *and* the base year was 74.3 percent. A total of 17 variables were used for the student

longitudinal nonresponse bias analysis. These 17 variables resulted in 66 comparisons (tests). Bias was detected for 33.3 percent of the 66 tests (table 59) implemented with the student longitudinal weight (W2W1STU). After applying the nonresponse adjustments, no bias was statistically significant in any of the 66 tests. A 79.0 percentage point reduction was also seen in the median absolute relative bias. The detailed analyses are shown in table F-2 within appendix F.

6.6.3 Nonresponse Bias Analyses of Parent Responses

6.6.3.1 First Follow-up Student-level Contextual Nonresponse Bias Analysis

The overall parental weighted response rate for the students randomly selected for the first follow-up parent subsample was 72.5 percent (table 47). Information on the nonresponding parents, however, was not available for either weight adjustment (section 6.4) or for the nonresponse bias analysis. Consequently, student and school characteristics used in the student-level nonresponse bias analysis were used for the student home-life contextual analyses.

In total, 17 variables were used for the student-level contextual nonresponse bias analysis, including characteristics known for the base-year schools where the students were first selected for the study. Again, these 17 variables resulted in 66 comparisons (tests). Bias was initially detected for 25.8 percent of the 66 tests (table 59) implemented with the first follow-up student home-life contextual weight (W2PARENT). After adjusting the weights, no tests were found to identify significant levels of bias. Also, the median relative bias was reduced by 89.7 percentage points. The detailed analyses are shown in table F-3 within appendix F.

6.6.3.2 Longitudinal Student-level Contextual Nonresponse Bias Analysis

The weighted response rate for the student first follow-up home-life contextual subsample was 72.5 percent (table 47). Accounting for students *and* parents in the subsample who responded to the base year *and* first follow-up, the weighted response rate was reduced by 8.3 percentage points or 64.2 percent (table 47). The student home-life contextual nonresponse bias analysis initially identified 37.9 percent of the 66 tests (table 59) as having significant levels of bias at the 0.05 level using the contextual longitudinal weight (W2W1PAR). After adjusting the weights, only 3 percent of the statistical tests produced significant results. Additionally, the median relative bias was reduced by 90.4 percentage points. The detailed analyses are shown in table F-4 within appendix F.

6.6.3.3 Nonresponse Bias Reduction in Student-level Contextual Responsive Design

A responsive design was implemented in the HSLS:09 first follow-up as one additional method for reducing nonresponse bias in the contextual information for students in the parent subsample (see section 4.4.5). Through the use of propensity models, the parent cases with low likelihood of response (i.e., low propensity) were identified and targeted for additional recruitment efforts.

To assess the effectiveness of the responsive design, a nonresponse bias analysis was conducted using frame variables to compare bias levels before and after the inclusion of the low-propensity cases (parents) who eventually responded. In other words, was the extra effort administered to convert the low-propensity cases to respondents effective? A successful reduction in bias would be identified if statistical tests with the frame variables showed a statistically significant estimated bias before, *but not after*, the inclusion of the low-propensity cases. Overall, approximately 42.4 percent of the categories initially showed an estimated bias that was statistically significant (table 60). Consequently, after the inclusion of low-propensity cases, 25.8 percent of the categories show estimated bias to be statistically significant. Details of the analyses summarized in table 60 are provided in table F-5 in appendix F.

Table 60. Summary statistics for unit nonresponse bias analyses in the HSLS:09 student-level contextual responsive design: 2012

	Significant bias tests[1]		Median absolute relative bias[2]		
Analytic weight	Percent without low-propensity cases	Percent with low-propensity cases	Percent without low-propensity cases	Percent with low-propensity cases	Percent relative change
Student home-life contextual responsive design[3]	42.4	25.8	4.3	4.2	-4.3

[1] Bias significantly different from zero at the 0.05 level of significance. Before and after are in reference to the inclusion of low-propensity cases targeting during the responsive design. A total of 66 statistical tests were performed using the W2PARENT weight. The number 66 was used as the basis for the reported percentages.
[2] The percent relative bias is calculated as 100 multiplied by the estimated bias divided by the estimated value. The absolute relative bias is the absolute value of the (percent) relative bias.
[3] Details of the responsive design associated with the student home-life contextual data are provided in section 4.4.5.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Data File.

6.7 Weighting Quality Control

A quality control (QC) phase was implemented for all activities, including the construction of the first follow-up weights. Various analytic properties of the initial weights, the weight adjustment factors, and the resulting weights after applying the adjustments were examined both overall and within sampling strata, including the (1) distribution of the weights, (2) ratio of the maximum weight divided by the minimum weight, and (3) unequal weighting effect. Finally, the sum of the weights were verified against pre-adjusted weight sums (e.g., marginal totals of the student weights prior to nonresponse adjustment and of respondents after nonresponse adjustment) and against the counts used in the final calibration adjustment. Similar procedures were used to QC the two sets of first follow-up BRR weights. As with the base year, a senior statistician also thoroughly checked each set of weights, owing to the central importance of these values in the calculation of population estimates.

To complement the standard set of QC weighting procedures, the design effect and unit nonresponse bias analyses were used. Relatively large design effects were examined to

determine whether variations in the adjustment factors were excessive, and bounds on the adjustments needed to be tightened. Results from the preliminary and final nonresponse bias analyses were examined to evaluate the effectiveness of the nonresponse model.

This page left intentionally blank.

Chapter 7. Item Response, Imputation, and Disclosure Treatment

7.1 Overview

Chapter 7 details the High School Longitudinal Study of 2009 (HSLS:09) first follow-up study procedures used to address patterns of response among those sample members who participated in the study and to protect the confidentiality of the HSLS:09 participants. Section 7.2 and appendix G contain the results from an analysis to evaluate detectable levels of bias associated with item nonresponse. Section 7.3 highlights the procedures and results associated with imputing a single missing value for a set of important study variables. Multiple imputation was also implemented on a few continuous variables as discussed in section 7.4. Additional information on the imputation methodologies implemented for the HSLS:09 first follow-up are included in appendix H. This chapter concludes in section 7.5 with a brief overview of methods used to evaluate the HSLS:09 data for disclosure risk and to treat the data to minimize the likelihood of identifying any particular sample member.

7.2 Item Nonresponse Bias Analysis

Item nonresponse bias, as with the unit nonresponse bias (section 6.6), affects the analytic results when those who should have provided a response but do not are different in some way relevant to the study from those who do provide a response. A description of the item nonresponse bias analysis conducted on the HSLS:09 first follow-up data is presented in section 7.2.1. The bias formula used is a function of the difference between the estimated values for item respondents and item nonrespondents, and (base) weighted item response rates among the eligible sample members.

Item response rates, in general, measure the proportion of responses obtained for a particular question among study respondents who were supposed to answer the question.[61] For example, if a student answers that he or she is not Hispanic, then the instrument routes around the subsequent Hispanic origin question. The value for the Hispanic origin variable is appropriately missing and is recoded to −7 in the HSLS:09 data file (see section 8.1). Conversely, if a student responds “yes” to the Hispanic question but does not provide their Hispanic origin, then the missing value for the latter question is labeled as item nonresponse. The weighted item response rate formula used in the nonresponse bias estimates is given in section 7.2.2.

[61] Item response rates differ from a unit response rate which measures the proportion of eligible sample members among those selected for the study who actually participate.

A weighted item response rate among study participants less than 85 percent, calculated with the final analytic weight as in the HSLS:09 base year, was used to identify first follow-up variables for the nonresponse bias analysis. The complete list of variables is provided in section 7.2.3. Finally, the item nonresponse bias results are summarized in section 7.2.4 and detailed in appendix G, section G.3.

7.2.1 Estimating Item Nonresponse Bias

The formula for estimating bias within HSLS:09 was first presented for unit nonresponse bias (section 6.7) among the set of eligible sample members selected for the study. An item-level analysis identifies detectable levels of item nonresponse bias specific to a certain variable within a given HSLS:09 study instrument among all eligible sample members.

The item nonresponse bias estimator has a similar form to the unit nonresponse bias estimator given in expression (6.20). Namely, item nonresponse bias is estimated as

$$\hat{Bias}\left(\overline{y}_{xR}\right) = \hat{\eta}_x\left(\overline{y}_{xR} - \overline{y}_{xNR}\right) \tag{7.1}$$

where x indicates the study item being analyzed for bias and $\hat{\eta}_x$ is the weighted item nonresponse rate among all eligible sample members calculated with the appropriate HSLS:09 base weight. Because item nonresponse negates the ability to calculate estimates for the item nonrespondents, the bias must be estimated using a characteristic y known for the item respondents *and* item nonrespondents. Here, the term "item nonrespondents" includes the set of unit respondents who were supposed to answer item-x but did not and the set of unit nonrespondents. Therefore, $\overline{y}_{xR}$ and $\overline{y}_{xNR}$ given in expression (7.1) are the estimated mean of y for the item respondents and nonrespondents, respectively. Note that the weighted nonresponse rate and the classification of unit respondents as item respondents or nonrespondents changes with each x-variable included in the analysis.

The y-variables for the item nonresponse bias analysis were chosen from a set of variables known for all sample members that were also associated with many important factors studied in HSLS:09. The following HSLS:09 first follow-up school characteristics were included in the analyses:[62]

- school type (public, private-total, private-Catholic, private-other);
- region of the United States (Northeast, Midwest, South, West); and
- locale (urban, suburban, town, rural).

[62] If school information was not available from the first follow-up data, base-year school characteristics were used in the analysis.

Student characteristics were also identified for the analyses:

- sex, and
- race/ethnicity (American Indian/Alaska Native, non-Hispanic; Asian, non-Hispanic; Black/African American, non-Hispanic; Hispanic, no race specified; Hispanic, race specified; More than one race, non-Hispanic; Native Hawaiian/Pacific Islander, non-Hispanic; White, non-Hispanic).

Prior to calculating the nonresponse bias estimates, the HSLS:09 data were edited for consistency and imputed values were excluded from the nonresponse bias analysis. Variables identified from the student survey were examined for potential nonresponse bias using both the first follow-up student analytic weight (W2STUDENT) and the student longitudinal weight (W2W1STU). Variables identified from the parent interview were examined for potential nonresponse bias using both the first follow-up student contextual analytic weight (W2PARENT) and the longitudinal weight for contextual analyses (W2W1PAR).

7.2.2 Item Response Rates

NCES statistical standards state that questionnaire items (or composite variables derived from a set of questionnaire items; see section 8.2 for details) with low item response should be examined for significant levels of nonresponse bias. This bias, as with unit nonresponse bias, could affect analysis results obtained from the study data and lead to erroneous conclusions.[63] All study items with a weighted response rate less than 85 percent among the study participants were classified as having high item-nonresponse and were included in the item nonresponse bias analyses.

Response rates for all HSLS:09 student and parent questionnaire items and composites were calculated as follows (see NCES Statistical Standard 1-3-5):

$$1-\hat{\eta}_x = \frac{I_x}{I-V_x}, \qquad (7.2)$$

the (weighted) number of sample members with a valid response to variable x (I_x) divided by the (weighted) total number of unit respondents (I) minus any cases for which the question was not applicable (V_x). The final analytic weights, adjusted for unit nonresponse and calibrated to population information, were used in the calculations.

The identification of the not applicable cases—study respondents who were excluded from the calculation—followed a specific set of rules. For example, if a student answered "no" to the following (gate) question on absence from school, then the subsequent set of questions on

[63] Nonresponse bias is defined as the difference between the estimated parameter calculated from the respondent data and the true value, and is estimated using weighted data from respondents and nonrespondents.

reasons for the absence would not be asked and the associated variables would have a not-applicable reserve code set.

Gate: *Has it been 4 or more weeks since you last attended high school?*

Branch: *Were you suspended or put on probation from the school?*

The value for the skipped questions would be coded as "−7" (= legitimate skip/not applicable) as described further in Chapter 8. All "−7" values were excluded from the item nonresponse bias analysis.

In contrast, if a question was not answered because the respondent (1) completed only a portion of the questionnaire or (2) completed an abbreviated questionnaire without the item after declining to complete the full instrument, then the respondent would be included as an item nonrespondent in the associated item nonresponse bias analysis.[64]

7.2.3 High Item-nonresponse Variables

7.2.3.1 Items from Student Questionnaire

A total of 36 items on or derived from the student questionnaire (7.1 percent unweighted of 504 items) were identified for the first follow-up nonresponse bias analysis using the W2STUDENT weight (table 61). The lowest weighted item response rate, 52.2 percent, was found for the "has taken IB math course(s)" question (S2IBMATH), a question administered to a small proportion of the students (2.5 percent). Almost 78 percent of the item-nonresponse bias analysis variables (28 of 36 items) had a weighted item response rate of at least 80 percent.

7.2.3.2 Items from Parent Questionnaire

A total of 33 items on the parent questionnaire (11.2 percent of 294 items) had a weighted response rate less than 85 percent using the parent weight W2PARENT (table 62). The lowest item response rate, 30.4 percent for the "Teenager helped respondent complete questionnaire" question (P2QHELP1), which was applicable to only about 4 percent of the cases. Over 54 percent of the item-nonresponse bias analysis variables (18 of 33 items) had a weighted item response rate of at least 60 percent.

[64] A total number of 222 first follow-up student respondents (1.1 percent unweighted) did not complete the entire questionnaire but provided sufficient information for analysis. Among the 8,651 first follow-up parent respondents, 262 (3.0 percent unweighted) did not complete the entire questionnaire and an additional 475 (5.5 percent unweighted) completed an abbreviated questionnaire. Note that sample members who were only ever administered the abbreviated questionnaire will have "−6" as values for the excluded items.

Table 61. Student-level questionnaire items with a weighted item response rate below 85 percent using W2STUDENT weight

		Percent of records by type of response[1]				
Variable name	Description	Valid	Not applicable	Item missing	Unweighted item response rate	Weighted item response rate[2]
S2MNOFAMREC	Not taking math because family member discouraged teen	9.9	88.5	1.7	85.5	85.0
S2MNOTCHRREC	Not taking math because teacher discouraged teen	9.9	88.5	1.6	85.8	85.0
S2MNOASSIGN	Not taking math because not assigned to it	9.9	88.5	1.7	85.5	85.0
S2MDONTDOWELL	Not taking math because doesn't do well in math	9.9	88.5	1.7	85.4	84.9
S2MNOCNSLREC	Not taking math because HS counselor discouraged teen	9.9	88.5	1.7	85.5	84.8
S2MNOPARREC	Not taking math because parent discouraged teen	9.9	88.5	1.7	85.4	84.8
S2MNOCAREER	Not taking math because won't be needed for career	9.9	88.5	1.7	85.4	84.8
S2MNOEMPREC	Not taking math because employer discouraged teen	9.9	88.5	1.7	85.3	84.7
S2MNOCLGSUCC	Not taking math because won't be needed to succeed in college	9.9	88.5	1.7	85.4	84.6
S2MNOFRIEND	Not taking math because friends were not taking it	9.8	88.5	1.7	85.1	84.4
S2SDISLIKE	Not taking science because really dislikes science	17.0	79.6	3.3	83.7	84.0
S2SNOTHSREQ	Not taking science because it is not required for HS graduation	17.0	79.6	3.4	83.3	83.6
S2STOOKBEFORE	Not taking science because took it earlier in the school year	16.9	79.6	3.5	83.0	83.3
S2SNOCNSLREC	Not taking science because HS counselor discouraged teen	16.9	79.6	3.5	82.9	83.2
S2SNOCAREER	Not taking science because won't be needed for career	16.8	79.6	3.5	82.8	83.2
S2SNOASSIGN	Not taking science because not assigned to it	16.9	79.6	3.5	82.9	83.2
S2SNOCLGADM	Not taking science because won't be needed to get into college	16.9	79.6	3.5	83.0	83.1
S2OCC30EARN	Expected earnings for choice of occupation at age 30	60.6	27.1	12.3	83.1	83.0
S2SNOFAMREC	Not taking science because family member discouraged teen	16.8	79.6	3.6	82.5	83.0
S2SNOTCHRREC	Not taking science because teacher discouraged teen	16.8	79.6	3.5	82.7	83.0

See notes at end of table.

Table 61. Student-level questionnaire items with a weighted item response rate below 85 percent using W2STUDENT weight—Continued

Variable name	Description	Percent of records by type of response[1]			Unweighted item response rate	Weighted item response rate[2]
		Valid	Not applicable	Item missing		
S2SNOPARREC	Not taking science because parent discouraged teen	16.8	79.6	3.5	82.7	83.0
S2SDONTDOWELL	Not taking science because doesn't do well in science	16.8	79.6	3.5	82.7	83.0
S2SNOEMPREC	Not taking science because employer discouraged teen	16.8	79.6	3.5	82.7	82.9
S2SNOCLGSUCC	Not taking science because won't be needed to succeed in college	16.9	79.6	3.5	82.8	82.8
S2SNOFRIEND	Not taking science because friends were not taking it	16.8	79.6	3.5	82.6	82.7
S2JOBYR	Year dropout/early grad started current/most recent job	0.8	99.1	0.2	82.7	82.5
X2PROBLEM[3]	X2 Scale of problems at high school	79.8	0.0	16.1	83.2	81.9
S2CHILDBORNYR	Year dropout/early grad's first child was born	0.6	99.3	0.1	82.4	80.9
X2S2EARNNOHS[3]	Earnings without HS diploma standardized by year	66.8	0.0	33.2	66.8	65.9
X2S2EARNHS[3]	Earnings with HS diploma standardized by year	66.6	0.0	33.4	66.6	65.8
X2S2EARN4Y[3]	Earnings with four year college degree standardized by year	65.2	0.0	34.8	65.2	63.9
X2S2EARNOCC[3]	Earnings with occupational training diploma standardized by year	64.5	0.0	35.5	64.5	63.5
X2S2EARN2YPUB[3]	Earnings with two year college degree standardized by year	64.5	0.0	35.5	64.5	63.4
S2IBOTHER	Has taken IB course(s) in another subject	2.5	95.6	1.9	57.4	52.8
S2IBSCIENCE	Has taken IB science course(s)	2.5	95.6	1.9	57.2	52.5
S2IBMATH	Has taken IB math course(s)	2.5	95.6	1.9	57.0	52.2

[1] The reserve codes "−7" and "−9" identify the legitimately skipped/not applicable questionnaire items and the questions that should have been answered but were not (item missing), respectively. The number of cases used in the calculation is the total number of eligible (unit) student respondents (n=20,594).

[2] Weighted response rates were calculated with the student analytic weight (W2STUDENT).

[3] Variable is a derived variable. Survey items used to create the derived variables are detailed in appendix I.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

Table 62. Parent-level questionnaire items with a weighted item response rate below 85 percent using W2PARENT weight

		Percent of records by type of response[1]				
Variable name	Description	Valid	Not applicable	Item missing	Unweighted item response rate	Weighted item response rate[2]
P2CLGWORKFT	Teenager will work full-time or part-time while attending college	24.0	72.0	4.0	85.7	84.7
P2CONF2YPUB	Confidence in estimate of cost of public in-state 2-year college	50.3	41.5	8.2	86.0	84.0
P2CONF4YPRV	Confidence in estimate for cost of typical 4-year private college	53.2	38.2	8.6	86.0	83.5
P2FIRSTCHOICE	Most likely postsec school is parent's 1st choice not considering cost	39.1	52.8	8.1	82.9	80.5
P2CERTAINCLG	How certain teenager is to attend most likely postsecondary institution	39.0	52.8	8.1	82.8	80.3
P2USYR1	Year Parent 1 came to U.S. to stay	19.0	75.6	5.5	77.7	76.1
P2USYR2	Year Parent 2 came to U.S. to stay	16.0	78.8	5.2	75.3	73.3
P2LIKELYCLGLV	Level of postsecondary institution most likely to attend in fall 2013	34.4	52.8	12.8	72.8	69.4
P2LIKELYCLGTYP	Control (public/private) of postsec inst most likely to attend in fall 2013	34.1	52.8	13.1	72.3	68.9
X2PEARNNOHS[3]	Parent questionnaire earnings without HS diploma standardized by year	67.3	0.0	32.7	67.3	66.2
X2PEARNHS[3]	Parent questionnaire earnings with HS diploma standardized by year	67.9	0.0	32.1	67.9	66.2
P2CHOICECLGLV	Level of parent's first choice postsecondary institution	29.1	57.4	13.5	68.3	65.7
P2CHOICECLGTYP	Control (public/private) of parent's first choice postsecondary institution	28.9	57.4	13.7	67.8	65.2
X2PEARN4Y[3]	Parent questionnaire earnings with four year college degree standardized by year	67.5	0.0	32.5	67.5	65.2
P2ENGLISH	English is regularly spoken in home	21.2	68.4	10.4	67.1	65.2
X2PEARN2YPUB[3]	Parent questionnaire earnings with two year college degree standardized by year	64.9	0.0	35.1	64.9	63.1
X2PEARNOCC[3]	Parent questionnaire earnings with occupational training diploma standardized by year	64.8	0.0	35.2	64.8	63.1
P2USGRADE	Grade level teenager was placed in when started school in U.S.	7.4	88.9	3.7	66.8	60.0
P2USYRT	Year teenager came to the U.S. to stay	7.4	88.8	3.8	66.2	59.4
P2DKHOWAPP	Won't apply for financial aid because does not know how	12.1	79.9	8.0	60.0	57.4

See notes at end of table.

Table 62. Parent-level questionnaire items with a weighted item response rate below 85 percent using W2PARENT weight —Continued

Variable name	Description	Percent of records by type of response[1]			Unweighted item response rate	Weighted item response rate[2]
		Valid	Not applicable	Item missing		
P2NODEBT	Won't apply for financial aid because family doesn't want debt	11.9	79.9	8.1	59.4	56.7
P2NOPLANS	Won't apply for financial aid because doesn't plan to continue education	11.9	79.9	8.2	59.0	56.5
P2INELIGIBLE	Won't apply for financial aid because may be ineligible/unqualified	11.9	79.9	8.2	59.4	56.4
P2FORMSDIFF	Won't apply for financial aid because forms are too difficult	11.8	79.9	8.2	58.9	55.9
P2CANAFFORD	Won't apply for financial aid because can afford college/school w/out it	11.9	79.9	8.1	59.5	55.1
P2NOQUALCRED	Won't qualify for financial aid because of credit score	10.7	80.8	8.5	55.7	53.3
P2NOQUALTEST	Won't qualify for financial aid because grades or test scores too low	10.6	80.8	8.6	55.4	52.8
P2NOQUALINC	Won't qualify for financial aid because income is too high	10.6	80.8	8.6	55.4	52.8
P2NOQUALFAM	Won't qualify for financial aid because family member didn't qualify	10.6	80.8	8.6	55.1	52.6
P2NOQUALPT	Won't qualify for financial aid because will attend part-time	10.4	80.8	8.8	54.1	52.0
P2QHELP2	Other family member helped respondent complete questionnaire	4.1	86.9	9.0	31.4	30.4
P2QHELP4	Someone else helped respondent complete questionnaire	4.1	86.9	9.0	31.4	30.4
P2QHELP1	Teenager helped respondent complete questionnaire	4.1	86.9	9.0	31.4	30.4

[1] The reserve codes "−7" and "−9" identify the legitimately skipped/not applicable questionnaire items and the questions that should have been answered but were not (item missing), respectively. The number of cases used in the calculation is the total number of (unit) respondents in the parent subsample (n=8,651).

[2] Weighted response rates were calculated with the school analysis weight (W2PARENT).

[3] Survey items used to create the derived variables are detailed in appendix I.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) First Follow-up.

7.2.4 Summarized Results for the First Follow-up

Nonresponse bias was evaluated for the items identified in the previous section as having low levels of item response by several important characteristics. Note that, unlike the analysis performed in the base-year documentation, unit nonrespondents were classified as item nonrespondents for this analysis. The detailed analysis tables are included in appendix G. The frequency distribution of the bias ratios (estimated bias divided by the standard error) by study instrument are summarized in table 63 where ratios larger than 2.0 suggest non-negligible levels of item nonresponse bias. For example, 38.6 percent of the 576 bias tests (= 36 variables crossed with 16 school and student characteristics) on the student questionnaire, analyzed using the student base weight, had a bias ratio greater than 2.0.

Table 63. Frequency distribution of the estimated bias ratios by study instrument

Study instrument	Analysis Weight	Range of bias ratio[1]	Frequency[2]	Percent[3]
Student	W2STUDENT	Total	576	100.0
		0 ≤ bias ratio < 2.0	354	61.5
		2.0 ≤ bias ratio < 5.0	209	36.3
		5.0 ≤ bias ratio	13	2.3
Parent	W2PARENT	Total	528	100.0
		0 ≤ bias ratio < 2.0	330	62.5
		2.0 ≤ bias ratio < 5.0	162	30.7
		5.0 ≤ bias ratio	36	6.8

[1] The bias ratio is calculated as the estimated item nonresponse bias divided by the estimated respondent value. The "total" row identifies the total number of calculations completed by study instrument.

[2] The number of calculations falling in the specified range of the bias ratio values. Unit nonrespondents were classified as item nonrespondents for this analysis. The student and parent base weights were used for the respective analyses.

[3] Unweighted percent of calculations falling in the specified range of the bias ratio values.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

The bias was evaluated for various characteristics and is summarized in tables 64 and 65. For example in table 64, 576 statistical tests (= 36 student items crossed with 16 school/student characteristics) for non-negligible item nonresponse bias in the student data were conducted. Approximately 52 percent of the tests were significant at the 0.05 level, indicating that the characteristics for the item respondents and nonrespondents were not statistically different for roughly half of the tests with unit nonrespondents classified as item nonrespondents. As shown in the next two columns, the overall average and median relative bias was small, suggesting that the level of bias across the 36 items and 16 characteristics may not be substantively meaningful. On average, the absolute relative bias (which ignores the positive and negative signs on the individual calculations) is less than 8 points and fluctuates depending on the characteristic used in the analysis. A similar interpretation is used for the parent items in table 65 where a summary of 528 statistical tests (= 33 parent items crossed with 16 school/student characteristics) using the parent base weight is given.

Table 64. Summary statistics for student-level item nonresponse bias analyses using W2STUDENT weight

Characteristics	Number of *t* tests	Percent[1] of significant *t* tests	Relative bias[2]		Absolute relative bias[3]	
			Average	Median	Average	Median
Total	576	52.1	-4.1	-1.3	13.9	7.6
School type						
Public	36	88.9	2.0	1.8	2.1	1.8
Private	36	88.9	-38.0	-36.3	39.9	36.3
Region						
Northeast	36	75.0	-26.1	-25.2	26.5	25.2
Midwest	36	41.7	-0.8	-1.8	5.6	5.1
South	36	36.1	10.3	2.5	10.4	2.5
West	36	47.2	0.3	-1.4	9.8	6.6
Locale						
City	36	80.6	-6.6	-12.5	15.1	12.6
Suburban	36	69.4	-9.1	-10.7	10.6	10.7
Town	36	75.0	9.5	18.4	20.5	18.6
Rural	36	41.7	8.5	13.0	15.1	13.3
Race/ethnicity						
Hispanic	36	25.0	-6.2	-6.8	10.3	7.5
Asian	36	83.3	-10.2	-21.8	34.4	28.5
Black	36	47.2	-3.2	-3.3	7.4	5.7
Other	36	22.2	3.3	3.8	4.8	4.0
Student sex						
Male	36	5.6	-3.8	-2.6	4.6	2.6
Female	36	5.6	4.4	3.0	5.3	3.0

[1] Unweighted percent of statistical tests with an item nonresponse bias significantly different from zero at the 0.05 significance level.

[2] The relative bias is calculated as the estimated bias divided by the estimated value.

[3] The absolute relative bias is the absolute value of the relative bias.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

Table 65. Summary statistics for parent-level item nonresponse bias analyses using W2PARENT weight

Characteristics	Number of *t* tests	Percent[1] of significant *t* tests	Relative bias[2]		Absolute relative bias[3]	
			Average	Median	Average	Median
Total	528	42.4	-0.7	-1.0	14.9	9.1
School type						
Public	33	48.5	-1.3	-0.9	1.6	1.0
Private	33	48.5	17.1	11.8	23.1	16.0
Region						
Northeast	33	39.4	-11.7	-13.1	12.3	13.1
Midwest	33	57.6	-3.5	4.2	12.9	8.6
South	33	30.3	0.0	-0.4	4.8	2.0
West	33	60.6	9.7	10.8	13.7	10.8
Locale						
City	33	54.5	4.3	-3.0	14.4	7.7
Suburban	33	6.1	-1.6	-3.6	5.6	5.6
Town	33	33.3	-0.8	5.8	13.5	10.3
Rural	33	30.3	-5.2	3.7	12.1	5.2
Race/ethnicity						
Hispanic	33	66.7	3.1	-10.8	26.1	20.1
Asian	33	45.5	9.5	-14.1	40.4	24.2
Black	33	48.5	-30.5	-30.3	30.5	30.3
Other	33	78.8	0.0	12.2	19.5	14.5
Student sex						
Male	33	15.2	1.6	1.7	3.7	2.6
Female	33	15.2	-1.7	-1.8	3.9	2.7

[1] Unweighted percent of statistical tests with an item nonresponse bias significantly different from zero at the 0.05 significance level.

[2] The relative bias is calculated as the estimated bias divided by the estimated value.

[3] The absolute relative bias is the absolute value of the relative bias.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

7.3 Single-value Item Imputation

Missing data in an otherwise complete study instrument occurs when a study respondent does not answer a particular question either intentionally (e.g., declined to answer a sensitive question) or unintentionally (e.g., missed one item within a set of related questions). Most statistical software packages exclude records that do not contain complete information. This is of great concern for multivariate analyses where a combination of missing values could greatly reduce the utility of the data file.

To alleviate the problem of missing data from a respondent record, statistical imputation methods were employed for the first follow-up similar to those used for the HSLS:09 base year.

Advantages of using imputed values include the ability to use all study respondent records in an analysis (complete-case analysis) which affords greater statistical power. Additionally, if the imputation procedure is effective (i.e., the imputed value is equal to [or close to] the true value), then the analysis results are possibly less biased than those produced with the incomplete data file.

A set of key analytic variables was identified for item imputation on data obtained from ninth-grade students as of fall 2009 who responded to the HSLS:09 first follow-up in 2012 and on responses gathered from parents of students in the first follow-up parent survey subsample. Values were assigned in place of missing responses through single-value imputation (or through derivation from imputed values) first for 4 variables from the student questionnaires (section 7.3.1), followed by 13 variables from parent questionnaires for students in the parent survey subsample (section 7.3.2). Indicator variables (flags) were included on the analysis file to allow users to easily identify the imputed values. The quality control and evaluative procedures are summarized in section 7.3.3.

7.3.1 Imputed Student Questionnaire Items

Four key analysis variables were identified for single-value imputation (table 66) from the edited HSLS:09 first follow-up data. Additional variables were considered for this list but were excluded because of either high item response rate or they were deemed to be of little analytic importance.

Table 66. Student questionnaire variables included in the single-value imputation by number and weighted percent of values imputed

Student questionnaire variables[1]	Number of values imputed[2]	Weighted percent imputed[3]	Method of imputation
Student is Hispanic (X2HISPANIC)	4	0.02	Statistical
Student's native language (X2NATIVELANG)	3	0.01	Statistical
Student's race (X2RACE)	4	0.02	Derived[4]
How far student expects to get in school (X2STUEDEXPCT)	32	0.18	Statistical

[1] The student variables were imputed sequentially and are listed in the order in which they were imputed or derived.

[2] The number of values imputed is the unweighted count of responding students with missing information among those where a valid response should have been provided and includes imputations from base year used in place of missing first follow-up information. The total number of responding students was 20,594 and excludes 251 first follow-up questionnaire-incapable students.

[3] The final student analytic weight (W2STUDENT) was used to calculate the weighted percent imputed among all first follow-up respondents (20,594).

[4] The variable was derived from a combination of source variables, some containing imputed values. The imputation flag for this variable was set to "yes" (= 1) if any corresponds with the flag(s) for the source variable(s).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

7.3.1.1 Imputation Methodology

The imputation methodology implemented to address the missing data items for the variables in table 66 varied by (1) the type of variable (e.g., categorical vs. continuous), (2) the relationship(s) between this variable and other HSLS:09 variables, and (3) the rate and pattern of missing values. This examination was implemented initially with draft study data and finalized only after all data were edited and re-examined.

Stochastic methods were used to impute the missing values. Specifically, a weighted sequential hot-deck (WSHD; statistical) imputation procedure (Cox 1980; Iannacchione 1982) using the final student analysis weight (W2STUDENT) was applied to the missing values for the variables in table 66 in the order in which they are listed. The WSHD procedure replaces missing data with valid data from a donor record (i.e., first follow-up student [item] respondent) within an imputation class. In general, variables with lower item nonresponse rates were imputed earlier in the process.

Imputation classes were identified using a recursive partitioning function created in R®.[65] With recursive partitioning (also known as a nonparametric classification tree or classification and regression tree [CART] analysis), the association of a set of questionnaire items and the variable requiring imputation are statistically tested (Breiman et al. 1984). The result is a set of imputation classes formed by the cross-classification of the questionnaire items that are most predictive of the variable in question. The pattern of missing items within the imputation classes is expected to occur randomly so that the WSHD procedure can be used. The input questionnaire items included the sampling frame variables and variables imputed earlier in the ordered sequence or were identified through skip patterns in the instrument and literature suggesting an association.

In addition to questionnaire items used to form the imputation classes, sorting variables were used within each class to increase the chance of obtaining a close match between donor and recipient. If more than one sorting variable was chosen, a serpentine sort was performed where the direction of the sort (ascending or descending) changed each time the value of a variable changed. The serpentine sort minimized the change in the student characteristics every time one of the variables changed its value.

Finally, analysis weights were used to ensure that the population estimate calculated with data including the imputed values (post-imputation) did not change significantly from the estimate calculated prior to imputation (pre-imputation). See, for example, the HOTDECK procedure in SUDAAN®.[66]

As an example, the variable "how far student expects to get in school" (X2STUEDEXPCT) in table 66 required item imputation for 32 responding students. A series of

[65] See www.r-project.org.

[66] See http://www.rti.org/sudaan/.

variables was included in the CART analysis for X2STUEDEXPCT (e.g., school sector [X2CONTROL], school region [X2REGION], student race/ethnicity such as Asian status [X2ASIAN]), student sex [X2SEX], student's base-year education expectation [X1EDUEX]), resulting in the identification of X2CONTROL and X2ASIAN as the statistically relevant variables to produce classes (i.e., explicit stratification) of sufficient size (e.g., 50 cases or greater with at least 50 percent [unweighted] item respondents) for imputation (see table H-1). Since X2SEX was also identified as important, but its inclusion would have created classes too small for imputation, the student's sex was used as a sorting variable within the classes formed by X2CONTROL crossed with X2ASIAN. The WSHD procedure was then implemented within the imputation classes (X2CONTROL by X2ASIAN) only after sorting the student respondent data within each class by X2SEX. Additional information on the WSHD methodology is found in, for example, Cox (1980) and Iannacchione (1982).

In general, groups of related variables that resulted in the same imputation classes were imputed en masse within a single WSHD routine. This was implemented to ensure consistency among the imputed values and proper levels of association exhibited in the item respondent data. These included, for example, variables created from a multiple-response item. Additional details on the imputation methodology are found in appendix H.

7.3.1.2 Imputation Results

Table 66 contains the order in which the student questionnaire variables were imputed in addition to the method of imputation used to resolve the missing data problems. At each step, several quality control procedures were used to maximize the utility of the imputed values. These are summarized in section 7.3.3 Additional details on the imputation methodology are found in appendix H.

7.3.2 Imputed Parent Questionnaire Items

A set of 13 variables (table 67) was chosen from the parent questionnaire for single-value imputation because of their analytic importance. Specifically, the parent survey items chosen for single-value imputation were related to the calculation of socioeconomic status (SES; section 7.4.2), imputed in the HSLS:09 base year, or calculated from values imputed in the first follow-up. As discussed for the student imputation, additional parent questionnaire variables were considered for missing data treatment but were excluded due to factors such as high levels of item response.

The imputation procedures for the parent-survey items were similar to those described for the student variables in the previous section. Specifically, a WSHD methodology was employed to randomly assign donor values from the set of parents responding to the item to those cases without a valid response. After each imputation, several quality control checks were administered to ensure consistency of the imputed values with responses obtained during the first

follow-up. These are summarized in section 7.3.3. Additional details on the imputation methodology are found in appendix H.

Table 67. Parent questionnaire variables included in the single-value imputation by number and weighted percent of values imputed for students in first follow-up parent survey subsample

Parent questionnaire variable[1]	Number of items imputed[2]	Weighted percent imputed[3]	Method of imputation
Parent 1 relationship to sample member (X2P1RELATION)	4	0.09	Statistical
Parent 2 relationship to sample member (X2P2RELATION)	23	0.25	Statistical
Parent 1 highest level of education (X2PAR1EDU)	223	2.94	Statistical
Parent 2 highest level of education (X2PAR2EDU)	189	2.68	Statistical
Highest level of education for parents (X2PAREDU)	257	3.48	Derived[4]
Parent 1 and 2 relationship pattern (X2PARPATTERN)	511	6.58	Derived[4]
Parent 1 current/most recent occupation: 2-digit O*NET code (X2PAR1OCC2)	445	5.61	Statistical
How far in school parent thinks sample member will get (X2PAREDEXPCT)	536	7.15	Statistical
Parent 2 current/most recent occupation: 2-digit O*NET code (X2PAR2OCC2)	389	4.93	Statistical
Number of 2012 household members (X2HHNUMBER)	928	11.12	Derived[4]
Parent 1 employment status (X2PAR1EMP)	810	10.32	Statistical
Total family income from all sources in 2011 (X2FAMINCOME)	774	9.54	Statistical
Parent 2 employment status (X2PAR2EMP)	644	8.02	Statistical

[1] The parent variables were imputed sequentially. The order in which the variables were imputed or derived by imputation group is provided in appendix H, tables H-2 through H-6. The order of the variables presented in this table is associated with group 1.

[2] The number of values imputed is the unweighted count of responding parents (for students in the parent survey subsample) with missing item information among those where a valid response should have been provided. A total of 11,952 students were selected for the first follow-up parent survey subsample and 8,651 parents responded.

[3] The final parent analytic weight (W2PARENT) was used to calculate the weighted percent imputed among those where a valid response should have been provided among the parent survey respondents (W2PARENT > 0). Those records where the question was not applicable (i.e., -7 values) were excluded from the results presented in the table unless imputed as not applicable.

[4] The variable was derived from a combination of source variables, some containing imputed values. The imputation flag for this variable was set to “yes” (= 1) if any corresponds with the flag(s) for the source variable(s). In addition to the variables listed in this table, several MOM and DAD variables were also derived from P1 and P2 variables if the relationship codes were biological, adoptive, or stepmother/stepfather. These additional variables inherited values from the corresponding P1/P2 variables themselves.

NOTE: O*NET = Occupational Information Network.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

The imputation methodology implemented to address the missing data items for the parent variables listed in table 67 varied by those factors noted for the student variables (section 7.3.1). Specifically, the methodology was adapted for (1) the type of variable (e.g., binary vs. categorical), (2) the relationships between this variable and other HSLS:09 variables from the base year and first follow-up, and (3) the rate and pattern of missing values. Three additional factors were also considered in determining the approach for parent variable imputation:

(4) parent response status in the base year and first follow-up, (5) whether or not the parent composition in the first follow-up was known to be the same as in the base year using parent responses provided in both studies, and (6) the stability exhibited in certain variables from base year to first follow-up (i.e., “no change”). Factors (4) – (6) led to the creation of five imputation groups:

Group 1: First follow-up responding parent – base-year responding parent (6,666 cases);

Group 2: First follow-up responding parent – no base-year parent data (1,985 cases);

Group 3: No first follow-up parent data – base-year responding parent (8,579 cases)[67];

Group 4: No first follow-up parent data – no base-year parent data – first follow-up responding student – base-year responding student (2,793 cases); and

Group 5: No first follow-up parent data – no base-year parent data – first follow-up responding student – no base-year student data (1,145 cases).

Groups 1 and 2 combined comprised the item nonrespondents within the parent responding subsample regardless of the first follow-up student response status. Groups 3 through 5 combined contained cases for all first follow-up responding students not selected for the parent survey subsample. Cases within each group are considered members of either the “donor set” or the “imputation set” depending on the variable requiring imputation.

Missing values for the 13 variables were imputed within each group starting with group 1 and ending with group 5. Respondent-provided values in the current and subsequent groups were used as donors for the group of cases undergoing imputation, along with the imputed values from the previous group in the sequence listed above. For example, imputation group-2 donors included item respondent values from groups 1 and 2 and the imputed values obtained for the group-1 imputation set. The imputation procedures, both common and unique to each group, are discussed below.

Group 1: First follow-up responding parent – base-year responding parent

Missing values for 13 HSLS:09 parent questionnaire variables were imputed for the responding parents in the first follow-up parent survey subsample using the W2PARENT weight. Table 67 contains the list of variables and the number of items imputed for *all* parent survey respondents,[68] The first follow-up respondent set was divided into two groups for imputation—base-year respondents and base-year nonrespondents—so that variables available to ensure consistency for the imputed first follow-up values were included in the CART analysis.

[67] Note that group 3 includes cases with base-year parent data but with no first follow-up parent responses because either (1) the case was selected for the parent survey and the parent did not respond, or (2) the case was not selected for the parent survey.
[68] Counts by imputation group 1 and 2 are combined and not reported separately as a measure to lower confidentiality risk.

Group 1 comprised only those 6,666 cases with parent responses in both the first follow-up and the base year. Hence, variables from both studies were available for CART analysis to create the WSHD imputation cells. Table H-2 in appendix H contains the resulting list of imputation-cell variables identified from CART. Donor values, chosen to replace the missing values, were obtained from item respondents within the group-1 set of cases with parent responses in the base year and first follow-up.

Once the WSHD imputation was completed, several quality control (QC) procedures were implemented on the data (section 7.3.3). For example, the percent agreement for each variable between the base-year and first follow-up values was determined among the subset of group-1 cases where the parent composition was known to be the same in both the base year and first follow-up. The percent agreement between base year and first follow-up for the imputed first follow-up variables was verified to be similar to those cases with first follow-up responses. After completing all of the QC procedures, these data were then included in the donor pool for group 2. More information on the imputation procedures for Group 1 by parent questionnaire variable identified for item imputation is provided in appendix H, table H-2.

Group 2: First follow-up responding parent – no base-year parent data

The complement to group 1—those parent survey respondent cases without parent data collected in the base year—comprised group 2 of size 1,985. The donor cases were randomly selected through the WSHD methodology with the W2PARENT weight from all group-1 cases (whether imputed or not) along with group-2 cases with complete data (i.e., those that did not require imputation for the variable being analyzed). A CART analysis was implemented to identify imputation cell variables. Unlike group 1, the group-2 CART analysis, by definition, excluded base-year parent variables. Consistent with group 1, the QC procedures were implemented on the data to verify the quality of the imputed values prior to including the groups 1 and 2 (imputation and donor cases) in the donor pool for group 3. More information on the imputation procedures for Group 2 by parent questionnaire variable identified for item imputation is provided in appendix H, table H-3.

Group 3: No first follow-up parent data – base-year responding parent

To support analyses of student characteristics in concert with key home-life contextual variables, such as the components of SES, additional imputations were produced. To achieve information on variables such as parent education, occupation, and family income for the full set of responding students, stochastic imputation procedures were implemented on the same set of parent-questionnaire variables identified for responding first follow-up parents (table 67) within the set of 12,517 cases with first follow-up student survey data and no corresponding parent data. Note that because no parent information was provided for these cases during the first follow-up, the imputation methodology was implemented using other available data sources as discussed below and produced results that should be used with caution. Because the available information

from other sources varied within the 12,517 set of cases, three groups were created and imputed in turn (groups 3 through 5 specified above).

Missing values for the cases with base-year parent data, group 3 (n=8,579), were imputed using the W2STUDENT weight. Responding students with both base-year and first follow-up parent questionnaire information—either obtained directly through the parent questionnaire or imputed in the previous steps (groups 1 and 2)—formed the set of group-3 donor values for those students with base-year but without a first follow-up completed parent questionnaire (i.e., group-3 imputation set). This enabled the use of valid changes (or stability) from base year to first follow-up exhibited in the donor data for the missing values. The base-year parent responses (available for both the donor and recipient sets), along with variables used in the WSHD procedures for the responding parents, were examined as candidate imputation-class variables. For example, candidate variables including the corresponding construct derived in the base year, e.g., X1P1RELATION and X1P2RELATION (relationship of student to parent 1 and 2, respectively, in the base year) were examined for the imputation of X2P1RELATION (relationship of student to parent 1 in the first follow-up).

Once imputation was completed, several quality control procedures were implemented on the data (section 7.3.3). As with group 1, the most important pattern investigated was the percent agreement for each variable between the base-year and first follow-up values. As a result, a significant difference was discovered that needed to be addressed—imputations for cases with first follow-up parent interviews showed less change between the base year and first follow-up values compared to cases without first follow-up parent responses (e.g., change in highest parent education). This difference was attributed to the large set of variables used to form imputation classes for students with a responding first follow-up parent, in comparison to students with limited or no parent information.

To address the excessive amount of change in the parent variables (not already exhibited in the data) and in keeping with the spirit of WSHD, approximately 37 percent of the group-3 records in the imputation set were randomly chosen to *donate* their base-year parent questionnaire information in place of the missing first follow-up parent questionnaire values for eight of the 13 items listed in table 67: *parents' relationship to the student* (X2P1RELATION, X2P2RELATION); *parents' highest level of education* (X2PAR1EDU, X2PAR2EDU); *parents' employment status* (X2PAR1EMP, X2PAR2EMP); and *parents' current/most recent occupation code* (X2PAR1OCC2, X2PAR2OCC2).[69] Thus, "no change" was imputed between the base year and first follow-up for these randomly chosen cases. In keeping with the spirit of WSHD where information is "borrowed" from the respondent set, a proportion of cases was randomly chosen from student records with both base-year and first follow-up contextual data. The pattern

[69] The derived parent variables for these cases—*highest level of education* (X2PAREDU), *parent relationship pattern* (X2PARPATTERN), and *number of household members in 2012* (X2HHNUMBER)—were calculated from the resulting values. The remaining parent variables in table 67—*total family income* (X2FAMINCOME) and *how far in school parent thinks sample member will go* (X2PAREDEXPCT)—were allowed to remain the same or differ through the WSHD procedure.

analysis, along with other quality control procedures, was conducted once again before declaring the end of the imputation task for group 3. The final result was that the percent change the eight variables among responding students with parent responses in both the base year and first follow-up closely matched the percent change for first follow-up responding students with base-year parent survey data but without first follow-up parent survey data. More information on the imputation procedures for Group 3 by parent questionnaire variable identified for item imputation is provided in appendix H, table H-4.

Group 4: No first follow-up parent data – no base-year parent data – first follow-up responding student – base-year responding student

Reliance on base-year and first follow-up student responses was needed for group 4 imputation because no parent data were available from either study. Although no base-year parent data were available for classification in the set of 2,793 group-4 cases, the base-year socioeconomic status (SES) variable was derived for all base-year responding students.[70] For each variable in table 67, a candidate set of items was evaluated to construct the imputation classes through a CART analysis. The resulting subset of variables was then cross-classified by the base-year SES quintile (X1SESQ5) to produce the final classification for each variable. Including the SES quintiles in such a way allowed for variation in the first follow-up imputed values, using the base-year SES quintile as common base-year information for donors and recipients. Donor cases included those included in groups 1 through 3 discussed above, whether or not the values had been imputed. The WSHD procedure was implemented with W2STUDENT. More information on the imputation procedures for Group 4 by parent questionnaire variable identified for item imputation is provided in appendix H, table H-5.

Group 5: No first follow-up parent data – no base-year parent data – first follow-up responding student – no base-year student data

The final set of 1,145 cases requiring imputation of first follow-up parent questionnaire data only contained first follow-up student responses. In as much as possible, these variables were used to form imputation classes to ensure the (random) transfer of valid associations between first follow-up student and parent variables from donor to recipient. For example, when imputing X2P1RELATION, the corresponding first follow-up student variable, S2PARREL1, was used as a candidate classification variable. A CART analysis was again used to form the imputation classes with particular attention paid to ensuring class sizes had adequate numbers of donor and recipients. More information on the imputation procedures for Group 5 by parent questionnaire variable identified for item imputation is provided in appendix H, table H-6.

One final comment on parent imputation. It is worth noting that because no parent responses were obtained for first follow-up sampled students included in groups 3 through 5,

---

[70] Details of the base-year HSLS:09 SES are found in Chapter 7 of the base-year documentation. Note that the derivation of SES included not only the calculation of a score from component information but also multiple imputation of the SES values directly. Similar measures were used to calculate SES in the first follow-up (see section 7.4.2 below).

other data sources (when available) such as student responses and sampling information were used to form the imputation cells. Unlike the imputations discussed for groups 1 and 2, the utility of this source information for groups 3 through 5 is considered to be limited. The scarcity of data to form imputation cells with viable donor values is most apparent in group 5—first follow-up responding students with no base-year participation and no parent data—discussed above. Therefore, estimates produced with the values imputed (or derived) for the 12,517 cases lacking first follow-up parent data should be interpreted with this context in mind.

7.3.3 Evaluation of the Imputed Values

After each value or set of values was imputed, a set of quality control checks was implemented to ensure the highest quality. For example, the number of times a donor was used in the WSHD was tracked. If it was determined that a single donor was used too often (e.g., the donor value was used in place of three or more missing values) in a single imputation pass, then the classification process was altered to ensure a realistic level of variation in the imputed responses. Additionally, the percent agreement for each variable between the base-year and first follow-up values was determined among the subset of cases where the parent responded in both the base year and first follow-up and the household composition was known to be the same (see section 7.3.2, group 3 discussion). The weighted percent distribution of the values before and after the imputation procedure was also compared, both within and across the imputation classes, to identify large areas of change—see table H-7 for the student variables, table H-8 for the first follow-up parent variables, and table H-9 for the parent variables linked to respondents in both the base year and first follow-up. Differences greater than 5 percent at the 0.05 significance level were flagged and examined to determine whether changes should be made to the imputation sort or class variables.

Multivariate consistency checks ensured that relationships between the imputation variables were maintained and that any special instructions for the imputation were implemented properly. For these checks, it was important to ensure that the imputation process did not create any new relationships that did not already exist in the item respondent data. For example, if the imputed value for parental employment status (X2PAR1EMP and X2PAR2EMP) was “never worked for pay,” then the parental occupation variable (X2PAR1OCC2 and X2PAR2OCC2, respectively) was coded as a legitimate skip for consistency.

Additional consistency checks were implemented to compare data between the base year and the first follow-up. This step confirmed that no impossible combination of responses (or changes in the information) was introduced as a result of the first follow-up imputations. For example, if the education level for parent 1 in the base year (X1PAR1EDU) was “Bachelor’s degree” and the same parent did not respond to the question in the first follow-up, an imputation of “High school diploma” or less for X2PAR1EDU would be not be consistent. An imputed education level of “Bachelor’s degree” or higher would be consistent as long as any change in the education level was seen in the respondent data.

In any of the aforementioned checks, if there was any evidence of substantial deviation from the weighted sums or any identified inconsistencies, the imputation process was revised and rerun.

7.4 Multiple Imputation

Model-based multiple imputation (MI) procedures were reserved for four continuous variables within the HSLS:09 first follow-up. Through MI, the variance associated with imputation methodology is accounted for in the precision of the estimates provided that appropriate software is used. Variables identified for this method are generally continuous in nature and have a sufficient item nonresponse or relevance to the study that warrants the capture of the additional variation. The MI procedure was implemented on four HSLS:09 base-year variables. Owing to their analytic importance, the first follow-up versions of these variables were subjected to multiple imputation—the first follow-up student ability estimate for mathematics (*theta*; X2TXMTH) and the associated standard error of measurement for *theta* (*sem*; X1TXMSEM) simultaneously, followed by two versions of socioeconomic status simultaneously (SES; X2SES and X2SES_U). Details of the MI procedures for *theta* and *sem* are discussed in section 7.4.1, followed by a discussion of the methods for X2SES and X2SES_U in section 7.4.2.

7.4.1 Theta and SEM

A high completion rate was attained for the HSLS:09 first follow-up mathematics assessment. Among the 20,594 students who responded to the questionnaire, 89.9 percent or 18,507 students answered a sufficient number of assessment items to calculate both *theta* and *sem*. A set of five imputed values was generated for the remaining 2,087 students using SAS® PROC MI. The Markov Chain Monte Carlo model option, which assumes the data are from a multivariate normal distribution, was used to estimate the (joint posterior) probability distribution of the two variables. Random draws from this distribution were taken to fill in the missing values. As implemented in the base year, this simultaneous imputation was used to best capture the association of *theta* with *sem*.

The candidate predictor variables for the MI model used to impute *theta* and *sem* were taken from a large list of variables such as sex, race/ethnicity, student language, student postsecondary aspirations, parental aspirations for student, family composition, parental occupation and education level, household income, school type, locale, census division, and *theta* and *sem* from the base-year study (if available). Some combination of these variables from the first and second round of HSLS:09 was used to maximize consistency between the base-year and first follow-up assessment values. Prior to running the MI model, variables from this set were eliminated if they were highly correlated with other candidate model variables to ensure convergence to a stable solution. However, as is standard practice, many covariates were used in

the MI model to maximize the coverage of variables that might be used in models constructed by education researchers.

The imputation tasks resulted in six variables each for *theta* and *sem*. Variables X2TXMTH1–X2TXMTH5 and X2TXMSEM1–X2TXMSEM5 were created to contain the five imputed values for *theta* and *sem*, respectively. The average of the five *theta* imputed values was calculated and assigned to the variable X2TXMTH; the X1TXMSEM contains the average for the imputed *sem* values. The *theta* and *sem* values calculated for the assessment respondents (i.e., those not requiring imputation) were saved to the X2TXMTH and X2TXSEM and replicated within X2TXMTH1–X2TXMTH5 and X2TXMSEM1–X2TXMSEM5, respectively. The imputation flag X2TXMATH_IM distinguishes the imputed from the non-imputed values. The average values, the MI values, and the analysis weights (section 6.2) can be used with a variety of software to estimate the population value (see, e.g., SUDAAN®; PROC MIANALYZE procedure in SAS®; IVEware[71]).

Additional values were generated from the complete set (calculated and imputed) *theta* and *sem* values. Seven mathematics proficiency probability scores (X2TXMPROF1–X2TXMPROF7) were calculated from the five imputed values. Four variables were constructed from the average of the imputed *theta* values: the mathematics item response theory (IRT)–estimated number right score (X2TXMSCR); the mathematics IRT-estimated number right score defined by the base-year and first follow-up framework (X2X1TXMSCR); the standardized *theta* score (X2TXMTSCOR); and the *theta* score categorized into quintiles (X2TXMQUINT).

7.4.2 Socioeconomic Status

Two socioeconomic status (SES) indices were developed for the HSLS:09 base-year study using a slight variant of definitions used in previous NCES secondary longitudinal studies. Both rely on component variables for parent education, parent occupation, and family income. These same definitions are used again in the HSLS:09 first follow-up and are described below.

The first index (X2SES) was calculated to most closely match the definition used in other studies such as ELS:2002. The second index (X2SES_U), a variant of X2SES, accounts for differences in the target population by the locale of the base-year school (X1LOCALE).[72] Analysts who desire to account for the relativity of SES within locale have two options: (1) use X2SES_U in a bivariate or multivariate analysis, or (2) use X2SES in a multivariate analysis that controls for locale. An analyst who wants to achieve results that are more strictly comparable with those of the prior NCES secondary longitudinal studies should use X2SES. Both indices are briefly discussed below.

---

[71] See http://www.isr.umich.edu/src/smp/ive/.

[72] Geographic location of the base-year school was used for the first follow-up SES calculations instead of the first follow-up information (X2LOCALE) to maintain consistency with the base-year construct and consistency within the first follow-up since some students were no longer attending school.

The HSLS:09 first follow-up SES indices were constructed as a function of five component variables obtained from the parent/guardian questionnaire:

1. the highest education among parents/guardians in the two-parent family of a responding student, or the education of the sole parent/guardian (X2PAR1EDU);
2. the education level of the other parent/guardian in the two-parent family (X2PAR2EDU);
3. the highest occupation prestige score among parents/guardians in the two-parent family of a responding student, or the prestige score of the sole parent/guardian (X2PAR1OCC2);
4. the occupation prestige score of the other parent/guardian in the two-parent family (X2PAR2OCC2); and
5. family income (X2FAMINCOME).

Estimated means and standard deviations for the five SES components were calculated in SUDAAN® with the final first follow-up student analytic weight (W2STUDENT) on a set of records including (1) 7,490 students in the parent survey subsample with a valid response to all five SES components;[73] and (2) 893 students in the parent survey subsample with a response to at least two of the five SES components and no more than three (single-value) imputed SES components (table 68).[74] Standard procedures dictated that the data were edited for consistency prior to calculating a composite variable.

Means and standard deviations calculated from 8,383 records were used to generate the first SES index (X2SES). Means and standard deviations calculated within school urbanicity (X1LOCALE) were used to generate the second SES index (X2SES_U). With these estimates, five z-scores were calculated (one per SES component) for each index by subtracting the mean from the component value and dividing by the standard deviation using the following formula:

$$z_{ik} = \frac{x_{ik} - \hat{\bar{x}}_i}{std\left(x_i\right)}$$

where i is the index for the SES component ($i = 1, \ldots, 5$); k is the index for the student records; x_{ik} is the value for the ith SES component for the kth student; $\hat{\bar{x}}_i$ is the weighted mean of the ith SES component using W2STUDENT; and $std\left(x_i\right)$ is the estimated population standard deviation

[73] In the case of a single-parent household, a "legitimate skip" is considered to be a valid response for X2PAR2EDU and X2PAR2OCC2.

[74] Available categorical information was used to impute missing categorical values through WSHD procedures. Records without such information were designated for multiple imputation where the association among various base year and the first follow-up responses could be utilized to predict missing SES.

for the *i*th SES component calculated with W1STUDENT. SES for the *k*th student was then calculated as the unweighted average of the survey-based z-scores,[75]

$$SES_k = \frac{1}{5}\left(\sum_{i=1}^{5} z_{ik}\right).$$

Table 68. Distribution of responding students by SES imputation methodology and parent survey response status

Status	SES imputation methodology	Parent response status	Responding students	
			Number	Percent[1]
Total responding students or parents			20,919	100.0
Selected for parent survey subsample		Total	10,079	48.2
	Calculated SES (total)		8,383	40.1
	5 questionnaire responses	Respondent	7,490	35.8
	1–3 imputed components[2]	Item nonrespondent	893	4.3
	Multiple imputation (total)		1,696	8.1
		Item nonrespondent	238	1.1
		Nonrespondent	1,458	7.0
Not selected for parent survey	Multiple imputation	Total	10,840	51.8

[1] The unweighted percentages are calculated with respect to the total number of eligible, responding students (20,594). Percentages may not sum to 100 because of rounding.

[2] Available categorical information was used to impute missing categorical values through WSHD procedures.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

Five imputed values were randomly generated simultaneously for the remaining 1,696 students in the parent survey subsample in which either a student responded or a parent responded (238 students with responding parents but insufficient parent item-level responses plus 1,458 responding students with a nonresponding parent – 8.1 percent unweighted) and 10,840 responding students (51.8 percent unweighted) not selected for the parent survey subsample (table 68). Values for the 12,536 records were produced through an MI model that included the base-year SES (X1SES) but was otherwise similar to the model used for *theta* and *sem*.

The benefits of MI versus WSHD, a single-value imputation procedure, were weighed prior to imputing the missing SES values for the 12,536 cases. As discussed in section 7.3.2, WSHD was used to fill in missing values for the categorical SES components—parent education,

[75] A factor analysis approach to generate SES with differential component weights was also considered but never actually implemented. Because of (1) the comparison from one round of HSLS:09 to the next, and (2) the lack of an established model to use in a confirmatory factor analysis, a simple average calculation was used for HSLS:09 SES as in previous NCES studies.

parent occupation, and family income. One approach would have been to derive SES from the fully-imputed parent components, resulting in a single imputed value for the continuous index (per definition). In general, however, MI is used for imputation of continuous variables where values are "donated" from an underlying (continuous) distribution specified through the MI model, not "donated" from actual SES values derived from the un-imputed responses as with WSHD. Additionally, unlike single-value imputation procedures, MI affords the ability to capture in the estimated standard errors the additional level of uncertainty of the imputation process through the random selection of more than one (imputed) value. For these reasons, MI was chosen for SES imputation. The WSHD imputed components were used in the MI model (appendix H) to ensure consistency between the parent variables and the indices.

At completion, a set of HSLS:09 values was generated for the two SES indices. Five MI values for the first index (similar to other NCES surveys) were included in the variables X2SES1–X2SES5. The variables X2SES and X2SESQ5 contain the average of the five values and the X2SES quintile, respectively. The corresponding set of variables for the index controlled for base-year school locale is X2SES1_U–X2SES5_U, X2SES_U, and X2SESQ5_U. As with *theta* and *sem*, the 8,383 calculated values for SES were replicated in the set of MI variables (i.e., X2SES is equivalent to X2SES1 through X2SES5). Imputation flags were constructed for SES variables to identify three groups of records:

- X2SES_IM = 0 (7,490 records requiring no imputation—all of the variables used to construct SES were provided in the first follow-up parent questionnaire);
- X2SES_IM = 2 (893 records where SES was constructed with one to three imputed components—parent was a respondent to the parent interview but one or more of the components used to derive SES was missing); and
- X2SES_IM = 3 (12,536 records requiring multiple imputation—the student or the parent responded to the study but: the parent respondent provided insufficient item-level responses; or the student responded but the parent was not selected for the parent survey or the parent was selected but did not respond).[76]

7.5 Disclosure Risk Analysis and Protections

Extensive confidentiality and data security procedures were employed for HSLS:09 first follow-up data collection and data processing activities. Data were prepared in accordance with NCES-approved disclosure avoidance plans. The data disclosure guidelines were designed to minimize the likelihood of identifying individuals on the file by matching outliers or other unique data from external data sources. Because of the paramount importance of protecting the confidentiality of NCES data that contain information about specific individuals, HSLS:09 first follow-up data files were subject to various procedures to minimize disclosure risk. The HSLS:09 first follow-up data products and some of the disclosure treatment methods employed

[76] No imputation flag for X2SES_U was required since the calculation of X2SES and X2SES_U only differ through the use of X1LOCALE.

to produce them are described in the following sections. Details have been suppressed from this document to maintain the desired level of confidentiality.

7.5.1 First Follow-up Data Products

Data produced for the HSLS:09 first follow-up include restricted-use data and public-use data. Both the restricted- and public-use data include a student-level file. The student files contain responses and associated derived variables from the HSLS:09 first follow-up student, parent, school administrator, and counselor survey instruments, as well as all variables included in the student-level base-year data file. Additional variables include those associated with survey-based analysis such as analysis strata and final analysis weights (see chapter 6).

The disclosure treatment developed for the HSLS:09 first follow-up comprised several steps:

- reviewed the collected data and identified items that may increase risk of disclosure;
- applied disclosure treatment to the high-risk items to lower the risk of disclosure;
- produced restricted-use data files that incorporate the disclosure-treated data; and
- produced public-use data files, constructed from the disclosure-treated restricted-use files, using additional disclosure limitation methods.

The disclosure treatment methods used to produce the HSLS:09 first follow-up data files include variable recoding, variable suppression, and swapping. These methods are described below.

7.5.2 Recoding, Suppression, and Swapping

The disclosure treatment methods used to produce the HSLS:09 first follow-up data files include variable recoding, suppressing, and swapping. Some variables that had values with extremely low frequencies were recoded to ensure that the recoded values occurred with a reasonable frequency. Other variables were recoded from continuous to categorical values. In this way, rare events or characteristics have been masked for certain variables.

Other variables were classified as high risk and were suppressed from the public-use file. The suppressing techniques included removing the response from the file (i.e., reset to a “suppressed” reserve code) or removing records entirely from the public-use file (e.g., student nonrespondents).

Swapping was applied to certain HSLS:09 first follow-up data items. All respondents were given a positive probability of being selected for swapping and swapping was carried out under specific targeted, but undisclosed, swap rates. In data swapping, the values of the variables being swapped are exchanged between carefully selected pairs of records: a target record and a donor record. By doing so, even if a tentative identification of an individual is made, uncertainty remains about the accuracy and interpretation of the match because every record had some undisclosed probability of having been swapped. Swapping variables were selected from all

questionnaires: parent, student, administrator, and counselor. Summary information for the treated HSLS:09 first follow-up variables through a comparison of the public and restricted-use files is included in appendix L.

Because perturbation (swapping) of the HSLS:09 first follow-up data could have changed the relationships between data items, an extensive data quality check was carried out to assess and limit the impact of swapping on these relationships. For example, a set of utility measures for a variety of variables was evaluated pre- and post-treatment to verify that the swapping did not greatly affect the associations. Also, if the analysis determined that the components of a composite variable should be swapped, then the composite variable was reconstructed after swapping.

However, in contrast to swapping, composite variables and their components could have been independently suppressed or recoded for inclusion in public-use files, resulting in a potential mismatch in the public-use file. In cases where recoding or suppression of composite variables and their components was carried out independently, public-use data users may not be able to recreate some of the composite variables provided in the public-use files. An example of this situation included variables where the response categories have been collapsed for disclosure protection. The corresponding composite variable was derived from the full set of response categories as collected. Therefore, users who recalculate the composite variable with public-use information may see different results.

This page left intentionally blank.

Chapter 8.
Data File Contents

This chapter provides a concise account of the High School Longitudinal Study of 2009 (HSLS:09) base-year to first follow-up longitudinal data file contents. It addresses the construction and contents of the data file including composite variables and analysis weights.

8.1 Base-year to First Follow-up Data File

HSLS:09 first follow-up data have been made available in public-use versions in an electronic codebook (ECB) format and via the web-based Education Data Analysis Tool (eDAT) available at http://nces.ed.gov/edat/, and (for licensed users) restricted-use versions[77] in an ECB format. The ECB is installed from a DVD and is designed to be run in a Windows environment. The ECB (public-use version: NCES 2014-358; restricted-use version: NCES 2014-359) is available at no cost from the National Center for Education Statistics (NCES).

The ECB system serves as an electronic version of a fully documented survey codebook. It allows the data user to browse through all HSLS:09 variables contained on the data files, search variable and value names for key words related to particular research questions, review the wording of these items along with notes and other pertinent information related to them, examine the definitions and logic used to develop composite and classification variables, and export SAS, SPSS, or Stata syntax programs for statistical analysis. The ECB also provides an electronic display of the distribution of counts and percentages for each variable in the dataset. Analysts can use the ECB to select or “tag” variables of interest, export codebooks that display the distributions of the tagged variables, and generate SAS, SPSS, and Stata program code (including variable and value labels) that can be used with the analyst’s own statistical software.

The ECB comprises two files, one at the student level and one at the school level. The student-level file encompasses:

- Base-year student-level weights
- First follow-up student-level weights
- Base-year student-level composites
- First follow-up student-level composites
- Base-year student questionnaire
- First follow-up student questionnaire
- Base-year parent questionnaire
- First follow-up parent questionnaire

[77] A license is required to access the restricted-use ECB (http://nces.ed.gov/statprog/confid6.asp).

- Base-year teacher questionnaires
- Base-year administrator questionnaire replicated at student level
- First follow-up administrator questionnaire replicated at student level
- Base-year counselor questionnaire replicated at student level
- First follow-up counselor questionnaire replicated at student level
- Taylor series PSU and Stratum identifiers
- Balanced repeated replication weights

Analysts should be aware that the base-year school data may be used as a standalone, nationally representative sample of 2009–10 schools with ninth grades; however, the school data collected in the first follow-up are not generalizable to the nation's high schools with eleventh grades and therefore are not available as a separate school file. First follow-up administrator and counselor questionnaires are available only at the student level as these data apply only to student-level analyses. The school-level file has not changed since the base-year ECB and encompasses:

- Base-year school-level composite variables and weights
- Base-year administrator questionnaire
- Base-year counselor questionnaire

Data users should find naming conventions for variables, composites, and weights intuitive. Variables begin with an indicator of the data source, followed by a wave indicator, and finished with a descriptive name that easily identifies the variable. Specifics for the two-character prefix are below.

The first character distinguishes among components:

- X—composite variables
- W—weights
- S—student questionnaire
- P—parent questionnaire
- A—administrator questionnaire
- C—counselor questionnaire
- M—math teacher questionnaire (base year only)
- N—science teacher questionnaire (base year only)

The second character distinguishes the wave:

- 1—base year
- 2—first follow-up

Variable labels begin with the same two-character variable name prefix; however, additional information is provided to link users to the facsimiles and flowcharts, which include the section of the questionnaire (e.g., A, B, C) followed by the sequential numbering within the section. Some items have multiple components within the sequential numbering scheme and the section number receives a letter indicator (e.g., A04A, A04B, and A04C). Appendix K provides a detailed listing of all variable names and labels, and appendix N provides a listing of the critical items.

When data are missing for a particular item, negative value reserve codes are used to indicate why the item is missing. The following reserve code scheme is used:

- -1: “Don’t know” represents within continuous variables respondents who indicated that they did not know the answer to the question.
- -4: “Item not administered: abbreviated interview” is filled for questions that were not administered because an abbreviated version of the questionnaire was administered (e.g., F1 parent PAPI).
- -5: “Suppressed” represents values that have been suppressed on the public-use data files for disclosure reasons.
- -6: “Component not applicable” is filled for all variables across the entire questionnaire when a component did not apply (e.g., parents not included in the F1 subsample).
- -7: “Item legitimate skip/NA” is filled for questions that are not answered because prior answers route the respondent elsewhere.
- -8: “Unit nonresponse” is filled for all variables across the entire questionnaire when a sample member did not respond to the questionnaire.
- -9: “Missing” is filled for questions that are not answered within the questionnaire when the routing suggests that they should have filled in a response.

8.2 Composite Variables

Composite variables—also called constructed, derived, or created variables—are usually generated with responses from two or more questionnaire items or from the recoding of a variable (typically for disclosure avoidance reasons). Some are copied from another source (e.g., a variable supplied in sampling or imported from an external database). Examples of composite variables include school variables (school sector, school locale, region of the country), math assessment scores (achievement quintile in math), demographic variables (sex, race, Hispanic ethnicity, and month and year of birth), and the socioeconomic variables. Composite variable descriptions can be found in appendix E.

Most of the composite variables can be used as classification variables or independent variables in data analysis. Many of the composites have undergone imputation to address any missing responses in an attempt to lower item nonresponse bias. Note, all imputed versions of

variables have been flagged and are available in composite variables that are named with an IM suffix.

8.3 Analysis Weights

HSLS:09 is designed to produce accurate estimates for the two populations: (1) regular public schools (including public charter schools) and private schools in the 50 United States and the District of Columbia providing instruction to students in both the ninth and eleventh grades in the fall of 2009; and (2) ninth-grade students enrolled at the study-eligible schools across the United States (i.e., a national design) in the fall of 2009. Researchers must select the most appropriate weight for generation of the estimates from among the set of nine HSLS:09 analytic weights—five base year weights, two first follow-up weights, and two longitudinal weights (table 69). A brief summary of the discussion included in section 6.2 is given below to help in choosing the appropriate weight from the HSLS:09 base-year to first follow-up longitudinal data file.

Table 69. Available HSLS:09 weight variables for use in an analysis by the associated population and time period of interest

Potential HSLS:09 weight variables[1]	Population		Time period	
	School	Student	Base year	First follow-up
W1SCHOOL	✓		✓	
W1STUDENT		✓	✓	
W1PARENT		✓	✓	
W1SCITCH		✓	✓	
W1MATHTCH		✓	✓	
W2STUDENT		✓		✓
W2PARENT		✓		✓
W2W1STU		✓	✓	✓
W2W1PAR		✓	✓	✓

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-up Public-use Data File.

8.3.1 School-level Analyses

The W1SCHOOL weights are the only weights appropriate for use in producing estimates for the HSLS:09 population of schools. Estimates are only associated with the target population of schools in the base-year time period of fall 2009.

8.3.2 Student-level Analyses

Choice of a weight for student-level analyses can be a complicated issue for some researchers. To help in choosing a weight, researchers should first think in terms of the particular time period of interest for the HSLS:09 population of students—base year, first follow-up, or

both (i.e., longitudinal change between the base year and first follow-up). Next, researchers should consider the magnitude of nonresponse with the records included in the analyses and the associated nonresponse adjustment(s) for each weight. Both issues are considered below in producing estimates for the student population.

Student-level analyses, base-year only. Researchers interested in analyzing data associated only with the base-year time period have four weights from which to choose. The W1STUDENT weights are used with univariate or multivariate analyses involving the base-year student data (questionnaire and/or mathematics assessment). If contextual information is included in the analysis, then the weight is chosen that most reflects the nonresponse adjustment detailed in chapter 6. For example, if a desired analytic model includes student and parent data, then the W1PARENT weights are most appropriate. Substituting parent data for science or mathematics teacher data in this model requires the use of the W1SCITCH or W1MATHTCH weights. The inclusion of multiple contextual data sources requires some additional considerations.

The choice among the three contextual weights should rest with the patterns of nonresponse when combinations of any or all of the three contextual data sources are desired in the analysis. Note that records are excluded from a regression model if model covariates are missing, if the analysis weight is zero, or both. Consider an example where both parent *and* science teacher data are desired for a regression model to produce base-year student-level estimates. Using the rules above, two weights may be appropriate, W1PARENT and W1SCITCH. Both weights account for nonresponse patterns in the respective contextual data sources (i.e., parent and science teacher nonresponse patterns, respectively). However, since neither address nonresponse from both parents *and* science teachers, the use of either weight will be less than optimal. If the records available for the regression model (i.e., contain non-missing covariates) have a higher number of positive weights within one set, then that set of weights should be used in the analysis. Those records subsequently dropped from the model because of zero weights have no biasing effect on the estimates if they represent a portion of the student population that is no different from the portion covered by the model. Researchers may have to consider a different model specification if such an assumption is not reasonable. Note that the inclusion of either the administrator or counselor data, or school-level characteristics does not change the recommendations discussed above.

Student-level analyses, first follow-up only. Researchers interested in analyzing data associated only with the first follow-up time period have two weights from which to choose. The W2STUDENT weights are used with univariate or multivariate analyses involving the first follow-up student data (questionnaire and/or mathematics assessment). If first follow-up parent data are included in the analyses, then the W2PARENT weights are most appropriate.

Student-level, longitudinal analyses. The W2W1STU weights are most appropriate for addressing research questions associated with changes in student data from the base year to the

first follow-up. The W2W1PAR weights are suitable for longitudinal models that include home-life characteristics collected with the base-year and first follow-up parent questionnaires. If other contextual information is desired for the model (e.g., base-year teacher responses), then the decision rests in whether or not the parent longitudinal data will also be included. For example, if a student-level longitudinal analysis is desired without controlling for home-life (longitudinal) changes, then the W2W1STU weights are most appropriate regardless of inclusion of base-year teacher data. If instead the home-life changes will be accounted for in the model, then the W2W1PAR weights are most appropriate regardless of inclusion of base-year teacher data.

A note on teacher contextual data. Having discussed the choice of weights for analyses, some background caveats must be entered about the design and statistical weighting of the teacher component of HSLS:09. Several additional elements of the study design speak to a need for caution in using the teacher data for longitudinal analysis: (1) mathematics achievement was measured at the beginning of ninth grade and the end of eleventh grade, but teacher characteristics were only measured for fall of ninth grade; (2) teachers were not asked to rate or comment upon the individual HSLS:09 student; (3) very little curricular or classroom-level information was collected; and (4) students were linked to courses as represented by course titles (e.g., Algebra II, or Geometry), but not to a specific classroom that met at a specific time and place (e.g., Algebra II, section 3, meeting at 9 a.m.). These caveats should be kept in mind when dealing with the base-year teacher data.

If base-year teacher data are used in conjunction with first follow-up data, the premise in selecting a weight as discussed above, applies. Consider an example where both first follow-up student data *and* base-year math teacher data are desired for a regression model to produce first follow-up student-level estimates. The likely weight for this analysis is W2STUDENT. This weight adjusts for the nonresponse patterns associated with first follow-up student data, but not for the nonresponse patterns associated with base-year math teacher data. Researchers are encouraged to examine the pattern of missing data associated with the base-year teacher component and the W2STUDENT weight. If such an analysis suggests that the data are not necessarily missing at random, a property important to low nonresponse bias, then experienced researchers may choose to investigate additional adjustments to the weights or to the data such as an appropriate imputation model. Note, however, that the public-use file has limited information for use in such adjustments. Consequently, any subsequent adjustment could introduce more bias, not less, than using the data and weights in their published state.

Additionally, as was stressed in the base year Data File Documentation, the teacher sample does *not* constitute a nationally representative (or school representative) sample of ninth-grade mathematics and science teachers. The two separate teacher samples (mathematics and science) were not independently selected but rather depend on a linkage to a sampled student who was selected for the study using probability methods and who both was enrolled in the requisite subject area and participated in the base year. Although it is possible to create teacher-level and course-level data sets using the base-year teacher data, they do not constitute valid

generalizable probability samples of teachers. For this reason, neither a teacher ID, nor statistical weights, have been provided to support this level of analysis. The teacher weights in the base year support use of teacher data only as an extension of the student record, with the student as the unit of analysis.

References

American Association for Public Opinion Research. (2011). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th edition. AAPOR. Available: http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions2&Template=/CM/ContentDisplay.cfm&ContentID=3156

Baker, R.P., Bradburn, N.M., and Johnson, R.A. (1995). Computer-assisted Personal Interviewing: An Experimental Evaluation of Data Quality and Cost. *Journal of Official Statistics, 11*(4), 413–431.

Binder, D.A. (1983). On the Variances of Asymptotically Normal Estimators From Complex Surveys. *International Statistical Review*, 51(3), 279-292.

Bock, D., and Mislevy, R.J. (1982). Adaptive EAP Estimation of Ability in a Microcomputer Environment. *Applied Psychological Measurement, 6*(4), 431-444.

Bozick, R., and Ingels, S.J. (2008). *Mathematics Coursetaking and Achievement at the End of High School: Evidence from the Education Longitudinal Study of 2002 (ELS:2002)*. (NCES 2008-319). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Breiman, L., Friedman, J. Stone, C.J., and Olshen, R.A. (1984). *Classification and Regression Trees*. New York: Chapman & Hall/CRC.

Breiman L., Friedman J. H., Olshen R. A., and Stone, C. J. (1984) Classification and Regression Trees. Wadsworth. R function = http://finzi.psych.upenn.edu/R/library/tree/html/tree.html.

Chromy, J.R. (1981). Variance Estimators for a Sequential Sample Selection Procedure. In D. Krewski, R. Platek, and J.N.K. Rao (Eds.), *Current Topics in Survey Sampling* (pp. 329–347). New York: Academic Press.

Cox, B.G. (1980). The Weighted Sequential Hot Deck Imputation Procedure. *Proceedings of the Section on Survey Research Methods* (pp. 721–726). The American Statistical Association.

Curtin, R., Presser, S., and Singer, E. (2000). The Effects of Response Rate Changes on the Index of Consumer Sentiment. *Public Opinion Quarterly 64*(4), 413–28.

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

Deville, J.C., and Särndal, C-E. (1992). Calibration Estimating in Survey Sampling. *Journal of the American Statistical Association*, *87*, 376–382.

Folsom, R.E., Potter, F.J., and Williams, S.R. (1987). Notes on a Composite Size Measure for Self-weighting Samples in Multiple Domains. In *Proceedings of the American Statistical Association, Section of Survey Research Methods*, 792–796.

Folsom, R.E., and Singh, A.C. (2000). The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and poststratification. *American Statistical Association*, Proceedings of the Survey Research Methods Section (ASA), 598–603.

Guarino, C.M., Hamilton, L.S., Lockwood, J.R., and Rathbun, A.H. (2006). *Teacher Qualifications, Instructional Practices, and Reading and Mathematics Gains of Kindergartners* (NCES 2006-031). U.S. Department of Education, Institute of Education Sciences. Washington, DC: National Center for Education Statistics.

Hambleton, R.K., and Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer.

Iannacchione, V.G. (1982). Weighted Sequential Hot Deck Imputation Macros. In *Proceedings of the Seventh Annual SAS Users Group International Conference*, 759–763.

Ingels, S.J., Dalton, B., Holder, T.E., Lauff, E., and Burns, L.J. (2011). *High School Longitudinal Study of 2009 (HSLS:09). A First Look at Fall 2009 Ninth-Graders*. (NCES 2011-327). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Ingels, S.J., Herget, D., Pratt, D.J., Dever, J., and Copello, E. (2010). *High School Longitudinal Study of 2009 (HSLS:09) Base-Year Field Test Report* (NCES 2010-01). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Ingels, S.J., Pratt, D.J. Herget, D.R., Burns, L.J., Dever, J.A., Ottem, R., Rogers, J.E., Jin, Y., and Leinwand, S. (2011). *High School Longitudinal Study of 2009 (HSLS:09). Base-Year Data File Documentation*. (NCES 2011-328). U.S. Department of Education, Institute of Education Sciences. Washington, DC: National Center for Education Statistics.

Ingels, S.J., Pratt, D.J., Jewell, D.M., Mattox, T., Dalton, B., Rosen, J., Lauff, E., and Hill, J. (2012). Education Longitudinal Study of 2002 (ELS:2002/12). *Third Follow-Up Field Test Report* (NCES 2012-03). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC.

Ingels, S.J., Pratt, D.J., Wilson, D., Burns, L.J., Currivan, D., Rogers, J.E., and Hubbard-Bednasz, S. (2007). *Education Longitudinal Study of 2002 (ELS:2002) Base-Year to Second Follow-up Data File Documentation* (NCES 2008-347). U.S. Department of Education, Institute of Education Sciences. Washington, DC: National Center for Education Statistics.

Keeter, S., Miller, C., Kohut, A., Groves, R.M., and Presser, S. (2000). Consequences of Reducing Nonresponse in a National Telephone Survey. *Public Opinion Quarterly, 64*, 125–148.

Keyfitz, N. (1951). Sampling with Probabilities Proportional to Size: Adjustment for Changes in the Probabilities. *Journal of the American Statistical Association*, *46*(253), 105–109.

Little, R.J.A., and Rubin, D.B. (2002). *Statistical Analysis with Missing Data*. New Jersey: John Wiley & Sons, Inc.

Liu, J. and Aragon, E. (2000). *Subsampling Strategies in Longitudinal Surveys*. Proceedings American Statistical Association.

LoGerfo, L.F., Christopher, E., and Flanagan, K.D. (2011). *High School Longitudinal Study of 2009 (HSLS:09). A First Look at Fall Ninth-Graders' Parents, Teachers, School Counselors and School Administrators.* (NCES 2011-355). U.S. Department of Education, Institute of Education Sciences. Washington, DC: National Center for Education Statistics.

Mantel, N., and Haenszel, W.M. (1959). Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *Journal of the National Cancer Institute*, 22, 710-748.

Merkle, D.M., and Edelman, M. (2009, March). An Experiment on Improving Response Rates and Its Unintended Impact on Survey Error. *Survey Practice*.

Peytchev, A., Baxter, R.K., and Carley-Baxter, L.R. (2009). Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error. *Public Opinion Quarterly, 73*(4), 785–806.

Pfeffermann, D. (1993). The role of sampling weights when modeling survey data. *International Statistical Review*, 61, 317–337.

Research Triangle Institute. (2008). *SUDAAN Language Manual, Release 10.0.* Research Triangle Park, NC: Research Triangle Institute.

Research Triangle Institute. (2012). *SUDAAN Language Manual, Release 11.0*. Research Triangle Park, NC: Research Triangle Institute.

Riccobono, J.A., Henderson, L.B., Burkheimer, G.J., Place, C., and Levinsohn, J.R. (1981). *National Longitudinal Study: Base Year (1972) through Fourth Follow-up (1979) Data File User's Manual.* U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Rock, D.A. (2007). *A Note on Gain Scores and Their Interpretation in Developmental Models Designed to Measure Change in the Early School Years*. (Report RR-07-08). Princeton, N.J.: Educational Testing Service. Available: http://www.ets.org/research/contact.html.

Rock, D.A. (2012). *Modeling Change in Large-Scale Longitudinal Studies of Educational Growth.* (Report ETS SPC-12-01). ETS R&D Scientific and Policy Contributions Series. Princeton, N.J.: Educational Testing Service. Available: http://www.ets.org/Media/Research/pdf/RR-12-04.pdf

Rock, D.A., and Pollack, J.M. (1991). *Psychometric Report for the NELS:88 Base Year Test Battery*. (NCES 91-468). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Seastrom, M. (2003). *NCES Statistical Standards* (NCES 2003-601). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Available: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003601.

Schouten, B., Cobben, F., and Bethlehem, J. G. (2009). Indicators for the Representativeness of Survey Response. *Survey Methodology, 35*(1), 101–113.

Stocking, M.L., & Lord, F.M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*, 201–210.

Tourangeau, R., Sebring, P., Campbell, B., Glusberg, M., Spencer, B.D., and Singleton, M. (1987). *The National Longitudinal Study of the High School Class of 1972 (NLS:72) Fifth Follow-up (1986) Data File User's Manual* (NCES 87-406c). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Valliant, R., Dever, J.A., and Kreuter, F. (2013). *Practical tools for designing and weighting survey samples*. New York: Springer.

Wolter, K. (2007). *Introduction to Variance Estimation*. Second Edition. New York: Springer-Verlag.

Woodruff, R.S. (1971). A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association, 66*, 411–414.

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

Zimowski, M., Muraki, E., Mislevy R., and Bock, D. (2003). *BILOG-MG*. Lincolnwood, IL: Scientific Software International.