High School Longitudinal Study of 2009 (HSLS:09) Base-Year to Second Follow-Up

Data File Documentation

JUNE 2018

Michael A. Duprey
Daniel J. Pratt
Donna M. Jewell
Melissa B. Cominole
Laura Burns Fritch
Ethan A. Ritchie
James E. Rogers
Jamie D. Wescott
David H. Wilson
RTI International

Elise M. Christopher
Project Officer
National Center for Education Statistics

NCES 2018-140
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
Betsy DeVos
Secretary

Institute of Education Sciences
Mark Schneider
Director

National Center for Education Statistics
James L. Woodworth
Commissioner

Sample Surveys Division
Chris Chapman
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, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

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

NCES, IES, U.S. Department of Education
Potomac Center Plaza (PCP)
550 12th Street, SW
Washington, DC 20202

June 2018

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

This report was prepared for the National Center for Education Statistics under Contract No. ED-IES-14-C-0112 with RTI International. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation
Duprey, M.A., Pratt, D.J., Jewell, D.M., Cominole, M.B., Fritch, L.B., Ritchie, E.A., Rogers, J.E., Wescott, J.D., Wilson, D.H. (2018). *High School Longitudinal Study of 2009 (HSLS:09) Base-Year to Second Follow-Up Data File Documentation* (NCES 2018-140). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved [date] from https://nces.ed.gov/pubsearch.

Content Contact
National Center for Education Statistics
(800) 677-6987
NCES.Info@ed.gov

Contents

PAGE

List of Tables vii

List of Figures ix

Chapter 1. Introduction 1
- 1.1 Overview of Data File Documentation (DFD) Report 1
- 1.2 Historical Background: NCES Secondary Longitudinal Studies Program 2
- 1.3 High School Longitudinal Study of 2009 5
  - 1.3.1 Base Year, First Follow-up, 2013 Update 6
  - 1.3.2 Second Follow-up 7
  - 1.3.3 Research and Policy Issues and Analytic Levels 9

Chapter 2. Sample Design 17
- 2.1 Base-year Sample Design 17
- 2.2 First Follow-up Sample Design 18
- 2.3 2013 Update and High School Transcript Study Sample Design 19
- 2.4 Second Follow-up Sample Design 20

Chapter 3. Instrumentation 23
- 3.1 Base-year, First Follow-up, and 2013 Update Instruments 23
- 3.2 Instrument Development in the Second Follow-up: Goals, Processes, Procedures 24
- 3.3 Survey Instrument Content in the Second Follow-up 25
- 3.4 Survey Instrument Design and Features in the Second Follow-up 29
- 3.5 Criteria for Defining Completed Interviews in the Second Follow-up 32

Chapter 4. Data Collection Methodology and Results 33
- 4.1 Data Collection Methodology 33
  - 4.1.1 Website 33
  - 4.1.2 Help Desk 35
  - 4.1.3 Data Collection Staff Training 35
  - 4.1.4 Tracing, Locating, and Interviewing Procedures 37
  - 4.1.5 Data Collection Quality Control Procedures 47
- 4.2 Responsive Design Methodology 48
  - 4.2.1 Data Collection Design Details 50
- 4.3 Data Collection Results 60

PAGE

4.3.1 Tracing and Locating Outcomes 63
4.3.2 Interview Participation Rates 66
4.3.3 Survey Timing 71

Chapter 5. Data Preparation and Processing 77
5.1 Overview of Systems 77
5.2 Data Cleaning and Editing 79
5.3 Coding 81
5.3.1 Major or Field of Study Coding 81
5.3.2 Occupation Coding 83
5.3.3 Respondent Job at Age 30 Coding 86
5.3.4 High School Coding 87
5.3.5 Postsecondary Institution Coding 88

Chapter 6. Response Rates, Analytic Weights, Variance and Design Effects Estimation, Nonresponse Bias Analysis, Imputation, and Disclosure Avoidance 91
6.1 Unit Response Rates 91
6.2 Overview of Weighting in the Base Year, First Follow-up, and 2013 Update 94
6.3 Second Follow-up and Supplemental 2013 Update Weights 95
6.3.1 Second Follow-up Weights 96
6.3.2 Supplemental 2013 Update Weights 101
6.3.3 Weighting Quality Control 103
6.4 Choosing an Analytic Weight 105
6.4.1 Base-year School-level Analysis 114
6.4.2 Base-year Student-level Analysis 114
6.4.3 First Follow-up Student-level Analysis 116
6.4.4 2013 Update and High School Transcript Student-level Analysis 118
6.4.5 Second Follow-up Student-level Analysis 123
6.5 Measures of Precision: Standard Errors and Design Effects 125
6.5.1 Standard Errors 125
6.5.2 Design Effects 129
6.6 Item-level Declined Response 133
6.7 Unit and Item Nonresponse Bias Analysis 136
6.7.1 Unit Nonresponse Bias Analysis 136
6.7.2 Item Nonresponse Bias Analysis 141
6.8 Single-value Item Imputation 151
6.8.1 Imputed Survey Questionnaire Items 152

PAGE

6.8.2 Evaluation of the Imputed Values 154
6.9 Disclosure Risk Analysis and Protections 155

Chapter 7. Data File Contents 157
7.1 Base-year to Second Follow-up Data File 157
7.1.1 Restricted-use Data 157
7.1.2 Public-use Data 158
7.1.3 Mixed-use Data Products 159
7.2 Contents of the Second Follow-up Data Products 159
7.3 Variable Naming Schema 162
7.4 Reserve Codes 163
7.5 Composite Variables 163
7.5.1 Revised SES indices 164

References 167

Appendices

A. Glossary of Terms
B. High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up Field Test Report
C. Cognitive Testing Results
D. Second Follow-up Survey Specifications
E. Notification Materials for Data Collection
F. Responsive Design Supplement
G. Unit and Item Nonresponse Bias Analysis
H. Second Follow-up and Supplemental 2013 Update Detailed Weighting Specifications ith Equations
I. Standard Errors and Design Effects
J. Imputation Details
K. ECB Variable Listing
L. Documentation for Composite Variables

This page intentionally left blank

List of Tables

PAGE

Table 1. Data collection staff trainings: 2016 36
Table 2. Hardcopy mailings: 2016 45
Table 3. Data collection schedule: 2016 50
Table 4. Calibration sample and main sample sizes, by subgroup: 2016 57
Table 5. Participation rates, by prior response status and student type: 2016 62
Table 6. Located status, by prior response status and student type: 2016 64
Table 7. Batch processing case match rates, by tracing source: 2016 65
Table 8. Intensive tracing rates and rates located though intensive tracing, by prior response status and student type: 2016 66
Table 9. Distribution of interview participation, by interview mode: 2016 67
Table 10. Participation status of field cases, by interview mode: 2016 68
Table 11. Distribution of interview respondents, by data collection phase: 2016 69
Table 12. Interview completeness among respondents: 2016 69
Table 13. Refusal and refusal conversion rates, by prior response status and student type: 2016 70
Table 14. Incentive selection among web and centralized telephone respondents who were offered an incentive: 2016 71
Table 15. Average time in minutes to complete the full interview, by interview section and mode: 2016 73
Table 16. Average time in minutes to complete the abbreviated interview, by interview section and mode: 2016 75
Table 17. Logically inferred variables: 2016 80
Table 18. Survey-uncoded majors, by coding method: 2016 82
Table 19. Results of quality control recoding and upcoding of major: 2016 83
Table 20. Results of quality control recoding and upcoding of occupation: 2016 85
Table 21. Job at age 30 results, by coding method: 2016 86
Table 22. Results of quality control of job at age 30 upcoding: 2016 87
Table 23. Final disposition of survey-uncoded high schools, after coding: 2016 88

PAGE

Table 24. Survey-uncoded postsecondary institutions, by coding method: 2016 89
Table 25. Final disposition of survey-uncoded postsecondary institutions, after coding: 2016 90
Table 26. HSLS:09 Base-weighted Unit Response Rates 93
Table 27. Descriptive characteristics of second follow-up survey weights 101
Table 28. Descriptive characteristics of supplemental 2013 Update survey weights 103
Table 29. HSLS:09 analytic weights 107
Table 30. Number and percentage of completed surveys, High School Transcript responses, or their combinations for the student sample, and associated recommended weights: Second follow-up 110
Table 31. Average design effects (*deff*) and root design effects (*deft*) for second follow-up student variables 132
Table 32. Item-level declined response over 5 percent, by variable and mode: 2016 134
Table 33. Summary statistics for unit nonresponse bias analyses before and after weight adjustments for nonresponse, by HSLS:09 second follow-up and supplemental 2013 Update analytic weights 138
Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT weight 144
Table 35. Frequency distribution of the estimated bias ratios 150
Table 36. Summary statistics for student-level item nonresponse bias analyses 151
Table 37. Student questionnaire and composite variables included in single value imputation by number and weighted percent of values missing 153
Table 38. Reserve code values: 2016 163

List of Figures

PAGE

Figure 1. Longitudinal design for the NCES Secondary Longitudinal Studies program: 1972–2025 4
Figure 2. Longitudinal design for the HSLS:09 9th-grade cohort: 2009–2025 6
Figure 3. HSLS:09 base-year student survey conceptual map 10
Figure 4. Student eligibility and fielding disposition from base year to the second follow-up: 2016 21
Figure 5. HSLS:09 second follow-up website: 2016 34
Figure 6. Tracing and locating procedures: 2016 39
Figure 7. Timeline of data collection phases and activities: 2016 40
Figure 8. Overall locating and interviewing results: 2016 63
Figure 9. Second follow-up weight construction process for the four weights with sequential adjustments for nonresponse 97
Figure 10. Example SAS-callable SUDAAN code to calculate an estimated mean and linearization standard error for a second follow-up student-level analysis 128
Figure 11. Example SUDAAN code to calculate an estimated mean and replicate (BRR) standard error for a second follow-up student-level longitudinal analysis 128
Figure 12. Example Stata code to calculate an estimated mean and linearization standard error for a second follow-up student-level analysis 128
Figure 13. Example Stata code to calculate an estimated mean and replicate (BRR) standard error for second follow-up student-level analysis 128
Figure 14. Example SAS code to calculate an estimated mean and linearization standard error for a second follow-up student-level analysis 129
Figure 15. Example SAS code to calculate an estimated mean and replicate (BRR) standard error for a second follow-up student-level longitudinal analysis 129

This page intentionally left blank

Chapter 1. Introduction

1.1 Overview of Data File Documentation (DFD) Report

This data file documentation report provides information and guidance for users of data from the base year through second follow-up of the High School Longitudinal Study of 2009 (HSLS:09), with a focus on the second follow-up data collection. 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.

This documentation is divided into seven chapters. Chapter 1 provides an introduction and outlines the organization of the documentation. It describes the historical background of HSLS:09 as part of the NCES secondary longitudinal studies program and supplies a study overview including levels of analysis and research questions. Chapter 1 also describes previous HSLS:09 data collections including surveys with students, parents, and various school personnel; the High School Transcript collection; and the survey and administrative data collections that comprise the second follow-up.

Chapter 2 describes the steps used to select the base-year sample and describes sampling through each subsequent follow-up, explaining the resulting sample for the second follow-up.

Chapter 3 provides summary information on base-year, first follow-up, and 2013 Update survey instruments followed by a more in-depth discussion of the second follow-up instrument.

Chapter 4 details the second follow-up data collection methodology and results, including data collection design, procedures, participation outcomes, and evaluations. This chapter also provides a description of the responsive design methodology used in the second follow-up and a review of the protocol's effectiveness.

Chapter 5 discusses the various data collection systems and system components used to conduct the second follow-up study and describes the data processing and post-collection editing.

Chapter 6 describes response rates, weighting, and other statistical procedures. This chapter presents information on response rates in the base year through the second follow-up. The chapter includes a section explaining the creation of the second follow-up weights and an overview of the impact of the weights on nonresponse bias. Also included are sections describing item and unit nonresponse bias analyses, imputation methodology, imputation results, and the disclosure avoidance procedures applied to the second follow-up data.

Chapter 7 describes the contents of the restricted- and public-use data files from the base year through the second follow-up. The chapter describes means by which data users can access the restricted- and public-use data, whether through electronic codebook (ECB), Online Codebook, or web-based PowerStats/QuickStats tools. The chapter also describes the composite variables created from the multiple data sources and analytic weight variables provided in the data files.

This documentation also contains the following appendixes:

A. Glossary of Terms
B. High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up Field Test Report
C. Cognitive Testing Results
D. Second Follow-up Survey Specifications
E. Notification Materials for Data Collection
F. Responsive Design Supplement
G. Unit and Item Nonresponse Bias Analysis
H. Second Follow-up and Supplemental 2013 Update Detailed Weighting Specifications with Equations
I. Standard Errors and Design Effects
J. Imputation Details
K. ECB Variable Listing
L. Documentation for Composite Variables

1.2 Historical Background: 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"[1] and the need for policy-relevant,

[1] Public Law 93-380, 93rd Congress, H.R. 69, August 21, 1974: An Act to Extend and Amend the Elementary and Secondary Act of 1965, and for Other Purposes. See https://www.congress.gov/bill/93rd-congress/house-bill/69 (Title V).

nationally representative longitudinal data on high 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 to examine the personal, familial, social, institutional, and cultural factors that may affect that development.

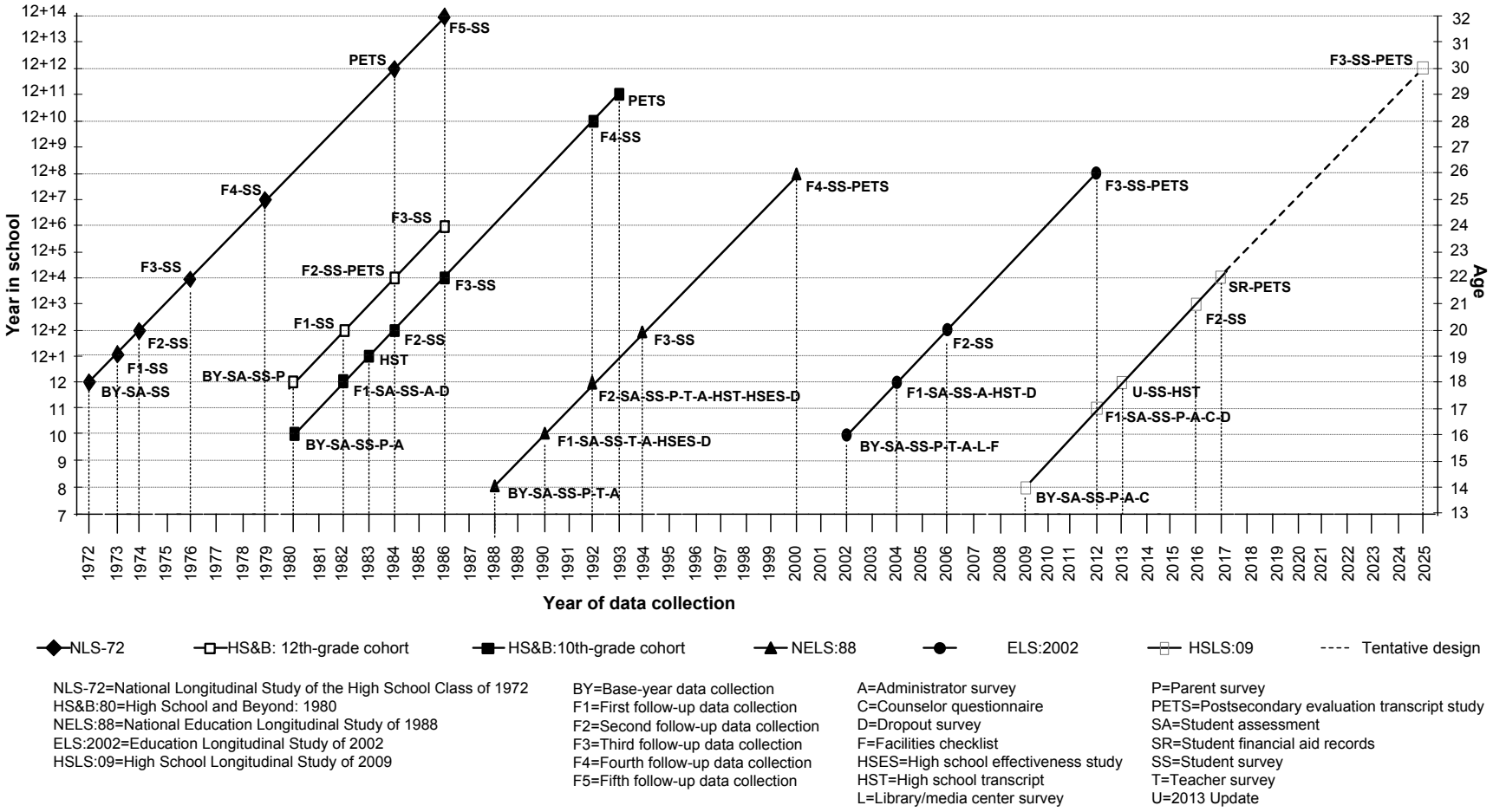
The Secondary Longitudinal Studies program consists of four completed studies, as well as the ongoing HSLS:09. The completed studies are the National Longitudinal Study of the High School Class of 1972 (NLS-72), the High School and Beyond Longitudinal Study of 1980 (HS&B:80), the National Education Longitudinal Study of 1988 (NELS:88), and the Education Longitudinal Study of 2002 (ELS:2002).

Together, these five studies describe the secondary and postsecondary experiences of students from five decades—the 1970s, 1980s, 1990s, 2000s, and 2010s—and also provide bases for further understanding the correlates of educational success in the United States. Information on both the current and completed studies in the series is available on the NCES website.[2]

Figure 1 presents a chronology of these five longitudinal education studies and highlights their component and comparison points for the time frame from 1972 to 2025.

[2] https://nces.ed.gov/surveys/slsp/

Figure 1. Longitudinal design for the NCES Secondary Longitudinal Studies program: 1972–2025



NLS-72=National Longitudinal Study of the High School Class of 1972
HS&B:80=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=First follow-up data collection
F2=Second follow-up data collection
F3=Third follow-up data collection
F4=Fourth follow-up data collection
F5=Fifth follow-up data collection

A=Administrator survey
C=Counselor questionnaire
D=Dropout survey
F=Facilities checklist
HSES=High school effectiveness study
HST=High school transcript
L=Library/media center survey

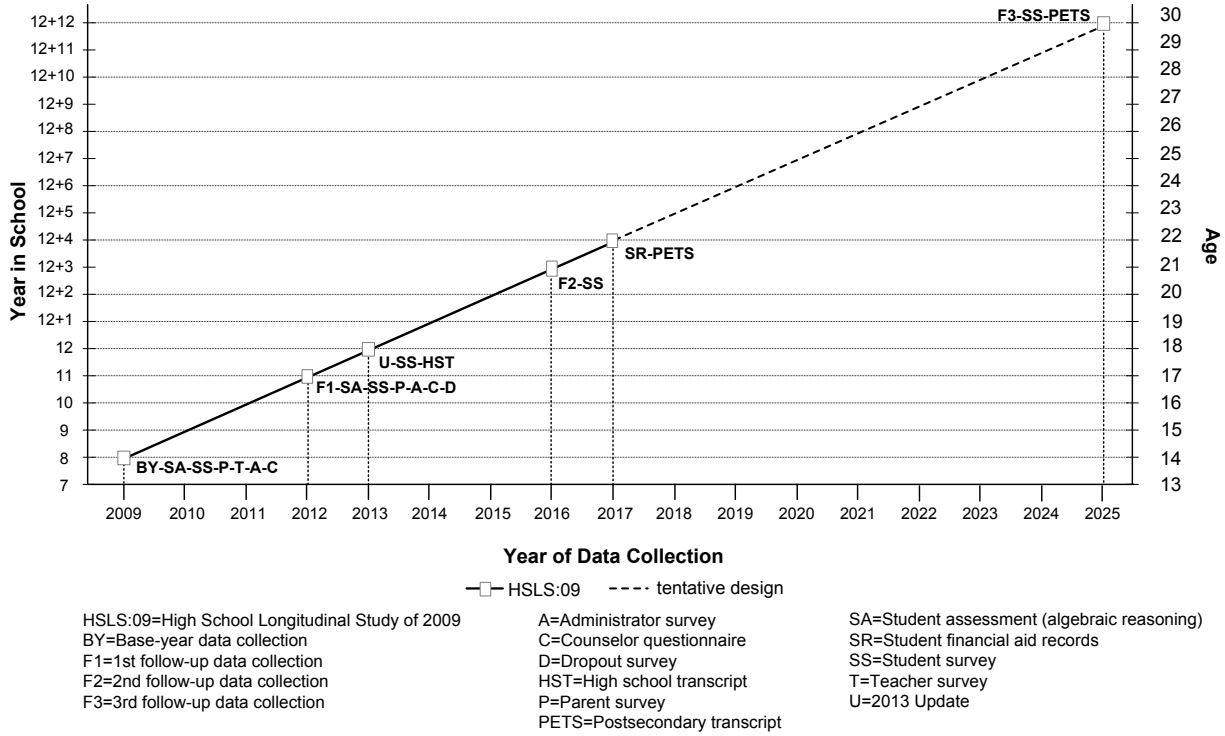
P=Parent survey
PETS=Postsecondary evaluation transcript study
SA=Student assessment
SR=Student financial aid records
SS=Student survey
T=Teacher survey
U=2013 Update

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

1.3 High School Longitudinal Study of 2009

HSLS:09 is based upon a nationally representative sample of entering freshmen in the fall of 2009 who were selected from a nationally representative sample of high schools with 9th and 11th grades. The study is designed to serve multiple policy objectives, primarily through longitudinal analysis. The goal of HSLS:09 is to provide data to understand better the impact of earlier educational experiences, starting at 9th-grade entry, on high school performance and the impact of these experiences on the transitions that students make from high school to adult roles. HSLS:09 was designed to help researchers and policy analysts investigate the process of dropping out of high school and possible return to school or pursuit of alternative credentials; the school experience and academic performance of English language learners; the nature of the paths into and out of science, technology, engineering, and mathematics (STEM) curricula and occupations; and the educational and social experiences that affect these outcomes, decisions, and experiences. The second follow-up extended the focus of the study to emphasize the transition of the cohort to postsecondary education—both baccalaureate and subbaccalaureate—and the workforce, including access to higher education and choice of postsecondary institution. The longitudinal design of HSLS:09 is illustrated in figure 2.

Figure 2. Longitudinal design for the HSLS:09 9th-grade cohort: 2009–2025



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

1.3.1 Base Year, First Follow-up, 2013 Update

The HSLS:09 base-year data collection took place in the 2009–10 academic year with a randomly selected sample of fall-term 9th-graders in more than 900 public and private high schools with both 9th and 11th grades.[3] Students completed an in-person mathematics assessment focused on algebraic reasoning and a web-based survey that included items on educational experiences, sociodemographic background, educational expectations, and their perceptions of the value of science and mathematics as a subject area and as a vocation. Students' parents, principals, and science and mathematics teachers, as well as their school's lead counselors, completed surveys on the phone or on the Web.

The first follow-up of HSLS:09 took place in the spring of 2012, when most sample members were in 11th grade. The students were again assessed in mathematics, and they again completed a questionnaire. The first follow-up survey explored topics such as high school attended, grade progression, school experiences, plans and

[3] Types of schools that were excluded from the sample based on the HSLS:09 eligibility definitions are described in the discussion of the target population in the *HSLS:09 Base-Year Data File Documentation* (see chapter 3, section 3.2.1) (Ingels et al. 2011).

preparations for the future transition out of high school, math and science aptitude and engagement, and extracurricular participation. Contextual data were again collected from a subsample of parents and from school administrators and counselors. While re-administration of the counselor questionnaire occurred only in the base-year schools, administrator questionnaires were administered at base-year schools as well as the schools to which students had transferred.

The 2013 Update occurred in the last half of 2013 (summer/fall of 2013). The 2013 Update was designed to collect information on the cohort's postsecondary plans and choices at the completion of high school (for most of the cohort). More specifically, information was collected about high school completion status, applications and acceptances to postsecondary institutions, education and work plans for the fall of 2013, financial aid applications and offers, choice of institution, and employment experiences. As part of the 2013 Update, high school transcripts were also collected in the 2013–14 academic year. Records matching (e.g., college admissions test scores, Free Application for Federal Student Aid [FAFSA] data, GED[4] data) also contributed to the dataset.

1.3.2 Second Follow-up

The second follow-up of the HSLS:09 cohort included a survey of sample members. In addition to the survey, the second follow-up included the collection of information from student financial aid records and postsecondary transcripts in 2017.

Second follow-up survey. The second follow-up survey was field tested from April through July 2015; the field-test report is appended (appendix B). The second follow-up main study data collection, conducted between March 2016 and January 2017, was designed to collect information from the cohort approximately 3 years after the modal high school completion date. At that time point, sample members may have been engaged in various activities, such as: enrollment in postsecondary education, employment, serving in the military, volunteering, interning or getting other job-related training, and starting a family. Some sample members may have only recently received, or may still have been working toward, a high school credential. The survey explored a variety of topics that include, but are not limited to, high school completion and experiences, enrollment history and future enrollment plans, employment and unemployment history, family and home-life characteristics, and personal characteristics (e.g., disabilities, sexual orientation and gender identity,

[4] The GED credential is a high school equivalency credential earned by passing the GED test, which is administered by GED Testing Service. See https://www.gedtestingservice.com/ged-testing-service for more information on the GED test and credential.

and civic engagement). The second follow-up survey also collected information on topics addressed in previous data collections, such as experiences, influences, and constraints on decision-making about postsecondary education, majors, and occupations with an emphasis on STEM fields.

Postsecondary transcript collection. The HSLS:09 transcript collection is the sixth in a series of Postsecondary Education Transcript Studies (PETS) of high school cohorts. The first (NLS:72) took place in 1984 and was followed by the HS&B sophomore cohort (1993), HS&B senior cohort (1986), NELS:88 (2000), and ELS:2002 (2013) PETS collections. Postsecondary education transcript studies have also been done in connection with the Beginning Postsecondary Students (BPS) and Baccalaureate and Beyond (B&B) longitudinal studies. A fundamental difference is that BPS and B&B are studies of students selected from a nationally representative sample of postsecondary institutions (NPSAS), while the high school studies are based on a grade-cohort-based secondary school sample. In addition, BPS captures all students entering postsecondary education for the first time, including students who begin their postsecondary education later in life, and the high school studies miss these late entrants if they begin their postsecondary education outside the study's time frame. Likewise, B&B is representative of baccalaureate recipients, and studies such as HSLS:09 and ELS:2002, which lack both late entrants and late completers, are not.

As an official institution record, the postsecondary transcript is a more reliable source of data regarding academic performance than is a student's self-report. The transcript collection for HSLS:09, designed similarly to that conducted for ELS:2002 and BPS:04/09, will provide much-needed information on the undergraduate experiences of 2009 ninth-graders who pursue postsecondary education in the years following high school. The combination of transcript data and other study data collected through interviews, by matching the sample to external data sources, and by student record collection will allow researchers to analyze paths taken by cohort members as they begin undergraduate education. Postsecondary transcripts provide a wealth of data on enrollment, including degree or certificate program, terms enrolled, dual enrollment status, course intensity when enrolled, and fields of study. Furthermore, transcripts provide coursetaking details, including subjects taken and credits and grades earned. These data provide important links among the sample members' secondary academic performance, plans and expectations, and pathways into the workforce.

Financial aid records collection. Previous secondary longitudinal studies collected student financial aid data from federal aid databases which detail federal aid exclusively. As such, a complete picture of all sources of student financial aid data,

including both federal aid and nonfederal aid, has been lacking in the secondary longitudinal studies, constituting a limitation in the utility of the study for analyses related to receipt of financial aid. Availability of financial aid is important at all points in the postsecondary process, including initial access and choice, persistence, transfer, and ultimate educational attainment. The financial aid data records collected from the institutions attended by HSLS:09 sample members will greatly increase the analytic utility of HSLS:09. Cumulative aid and debt can be calculated with scholarship, fellowship, grant, and loan amount data. The financial aid record collection will also yield detailed information about students' enrollment patterns, degree or program of study and progress toward degree, and costs of attendance.

1.3.3 Research and Policy Issues and Analytic Levels

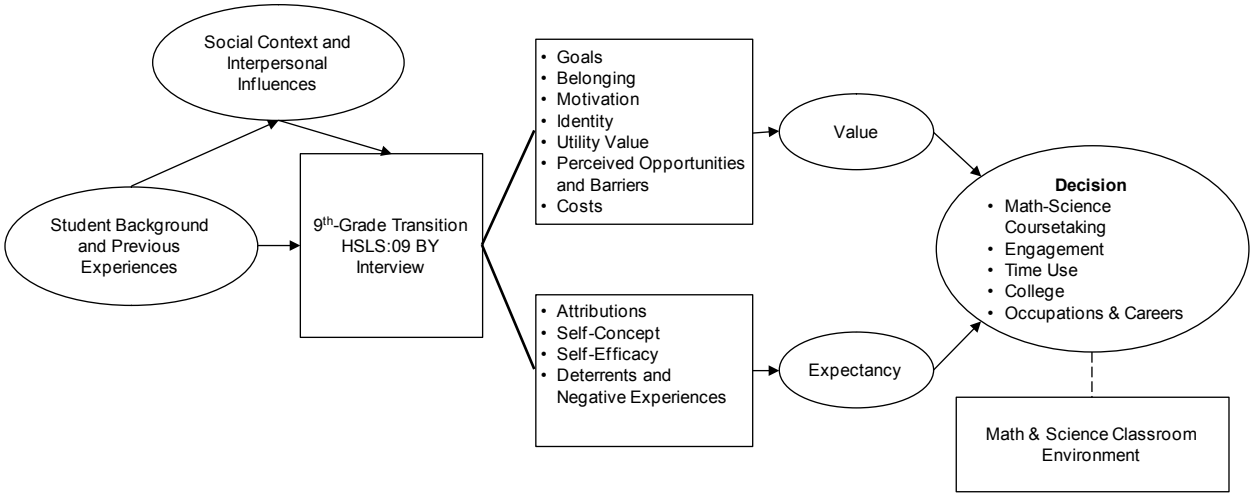
HSLS:09 is a general-purpose dataset, designed to serve multiple policy objectives rather than to test a specific hypothesis. The goal of HSLS:09 is better understanding of the relationship between earlier educational experiences, starting at 9th grade, and high school performance and the relationship of these experiences with the transitions that students make from high school to adult roles. HSLS:09 will help researchers and policy analysts investigate the features of effective high schools; growth in academic achievement, especially in mathematics;[5] factors related to dropping out of school and possible return to school or pursuit of alternative credentials; the school experience and academic performance of English language learners; the nature of the paths into and out of STEM curricula and occupations; and the educational and social experiences that affect these outcomes, decisions, and experiences.

The research agenda was guided by a conceptual model that was developed in the base year and shaped questionnaire content in both in-school rounds (i.e., fall 2009 base year and spring 2012 first follow-up). This model (figure 3) uses the student as the fundamental unit of analysis and attempts to identify factors that lead to academic goal setting and decision making. It traces the many influences, including motivation, interests, perceived opportunities, barriers, and costs, on students' values and expectations that factor into their most basic education-related choices. Results from the mathematics assessment, including achievement gains in the first 2 1/2 years of high school, are an important educational outcome measured in the study. Mathematics results can also predict readiness to proceed into postsecondary education, STEM courses, and careers. The study design also reflects the interaction between students and their families, with information collected directly from

[5] HSLS:09 includes an assessment of algebraic reasoning that measures achievement growth in the span between high school entry in the fall of 9th grade and the spring term of the junior year of high school for most cohort members (i.e., those in modal grade progression).

students' parents (or guardians) in the base year and first follow-up. It also captures the perspective of school administrators and counselors on the learning environment of the school. For the base year only, data are also available from 9th-graders' mathematics and science teachers.[6]

Figure 3. HSLS:09 base-year student survey conceptual map



The 2013 Update and the second follow-up built on the information collected during the base-year and first follow-up collections. The 2013 Update collected information on the cohort's postsecondary plans and choices, gathered at, for most of the cohort, completion of high school. More specifically, information was elicited concerning high school completion status, applications, and acceptances to postsecondary institutions, education and work plans for fall 2013, financial aid applications and offers, choice of institution, and employment experiences. The second follow-up was designed to collect a breadth of information on the cohort's pursuit of postsecondary education, entry into the workforce, and family formation. Furthermore, the addition of high school and postsecondary academic transcript information provides a continuous longitudinal record of courses taken, credit accrual, and grades in the high school though postsecondary years.

[6] The purpose of the HSLS:09 teacher surveys is to capture information on teachers' backgrounds, attitudes, and perceptions of the school climate. Information of this kind may contribute to the understanding of how teachers may encourage or discourage students in following the path to STEM and college. Teacher data were collected in fall 2009. Teachers were not asked about individual students the teacher taught, given the brevity of teacher-student exposure so early in the academic year.

In total, the breadth of the study design supports researchers in exploring a multitude of analytic interests and policy issues. Several examples of lines of investigation are outlined below.

Research and policy uses: base year and first follow-up. Many topic areas can be investigated within the high school context. These areas include the process of dropping out or stopping out of high school (e.g., taking a temporary break), the resilience of students who persist despite multiple risk factors, the educational and occupational trajectories of students who remain in school but take extra time to graduate, achievement gains in mathematics and the correlates of academic growth, the role of family background and the home education support system in fostering students' educational success, the features of effective schools, and differential access to and engagement in various educational opportunities.

Research and policy uses: 2013 Update. The 2013 Update was administered in the last half of 2013. For students who graduated on time, the timing of the data collection corresponded to collection immediately after completion of secondary school. The 2013 Update questionnaire consisted of objective questions that could validly be completed either by parent or student; there was no preference for which respondent should complete the relatively brief survey. It was designed to elicit critical time-sensitive data about how students and their parents make decisions about postsecondary choices. The 2013 Update provided information about status in summer/fall after the normative high school graduation, including educational status (e.g., high school completion, continued high school enrollment, high school dropout status, and postsecondary attendance); work status; postsecondary education applications and financial aid; and work experiences.

Research and policy uses: high school transcripts. Data from the HSLS:09 High School Transcript component encompass coursetaking for grades 9–12. Transcript data files can be analyzed on their own as stand-alone restricted-use files and can also be combined with the survey and assessment data for analysis.

High school transcript data from the secondary longitudinal studies also may be linked to postsecondary transcripts for high school cohort members who went on to postsecondary education or who were enrolled concurrently in postsecondary courses while in high school, known as "dual enrollment," thus providing information for analyses relating academic preparation and experiences in high school to coursetaking and attainment in higher education (Adelman 2006). At the high school level, evidence from HS&B (Cool and Keith 1991; Meyer 1998), NELS:88 (Rock and Pollack 1995), ELS:2002 (Bozick and Ingels 2008), and the National Assessment of Educational Progress (NAEP) (Chaney, Burgdorf, and

Atash 1997) suggests strong relationships between mathematics achievement and higher-level coursetaking.

Research and policy uses: second follow-up survey. Because most sample members in the 2016 second follow-up were 3 years beyond high school graduation, it is possible to study such topics as postsecondary education, entry into the workforce, and family formation.

The chief education-related foci of the second follow-up were access to postsecondary education, choice of postsecondary institution, and attainment of subbaccalaureate credentials. Early persistence and transfer from one postsecondary institution to another can also be studied. These topics of focus are asked of students who differ by postsecondary institution type and sector (e.g., public and private 2-year and 4-year institutions) attended; intensity of attendance (e.g., full-time versus part-time); whether enrollment was at the "first-choice" institution; and the institution's location (e.g., urban, suburban, or rural; near home or distant). A student's choice of postsecondary institution reflects institutional characteristics such as perceived academic quality or reputation, cost of attendance, and academic program offerings—all of which were captured in the 2013 Update and the second follow-up. The timing of the second follow-up also offered a window into attainment of 2-year degrees, postsecondary certificates, and certifications, whether granted by public institutions such as community colleges or by for-profit schools. The timing also provided an opportunity to view the transition from community college settings to 4-year programs for those sample members whose pathway treats 2-year institutions as a stepping stone to 4-year institutions. Other topics that can be explored include family formation; early occupational choice, with an emphasis on STEM fields; and labor market experiences.

Research and policy uses: postsecondary transcripts and student records. In addition to information obtained from sample members who participated in the second follow-up survey, data have been obtained as part of the student financial aid records and postsecondary transcript collection in 2017 from file matching to external sources and from institutions. Financial aid data records were collected from the institutions attended by HSLS:09 sample members, and federal student loan records were obtained from file matching. The financial aid data records collected from the institutions attended by HSLS:09 sample members will greatly increase the analytic utility of HSLS:09 and will yield detailed information about students' enrollment patterns, degree or program of study, progress toward degree, and cost of attendance. The postsecondary transcript data will cover postsecondary coursetaking through December 31, 2016 and provide detailed information on students' academic

experience, including coursetaking, academic performance, credit accumulation, enrollment periods, and transfer between institutions.

As mentioned above, data from high school transcripts from secondary longitudinal studies may be linked to data from postsecondary transcripts for high school cohort members who enrolled in postsecondary education or who were dual enrolled in postsecondary courses while still in high school, thus providing information for analyses relating academic preparation in high school to coursetaking and attainment in postsecondary education (Adelman 2006).

Research and policy uses: summary. HSLS:09 helps researchers, educators, and policymakers understand outcomes associated with the 9th-grade cohort's continued academic, social, and interpersonal status and growth in and after high school. It illuminates the transitions from postsecondary education to the workforce. It also captures students' choices about access to and persistence in STEM courses and majors, or alternative (i.e., non-STEM) educational and career pathways. Finally, it helps identify and describe the characteristics of educational institutions and curricula that are related to student outcomes in adulthood, such as family formation (e.g., how prior experiences in and out of school relate to marital or parental status and how marital or parental status affects educational choice, persistence, and attainment); and characteristics of individual students associated with key outcomes, including how language-minority, low-socioeconomic status (SES), disability, racial/ethnic-minority, and at-risk status are associated with education and labor market outcomes for young adults.

Analysis levels and design considerations. The base-year HSLS:09 data can be analyzed cross-sectionally at both the student and the school levels. Fall 2009 entering high school freshmen can be descriptively profiled using the HSLS:09 nationally representative student sample. Analysis at the school level is also possible, supported by the HSLS:09 nationally representative sample of high schools with 9th and 11th grades.[7] HSLS:09 obtained information about the base-year schools from several sources: a school administrator questionnaire; school characteristics variables taken from the sampling frame consisting of the NCES Common Core of Data (CCD) and Private School Universe Survey (PSS); and the school's course offerings, as listed in school catalogs collected in the High School Transcript study.

In addition to the national samples of high schools and fall 2009 9th-graders, the data support analysis of 10 state representative samples: California, Florida, Georgia, Michigan, North Carolina, Ohio, Pennsylvania, Tennessee, Texas, and Washington.

[7] Researchers should note that, due to disclosure protections, relatively few school-level analyses can be done with the public-use files; for most purposes, the restricted-use files are required.

The state samples pertain to the public sector only, and the national sample includes Catholic and other private schools.

Beyond the base year, HSLS:09 is representative of fall 2009 entering high school freshmen, who were followed up with 2 years after their 9th-grade year (first follow-up), the summer after the majority finished high school (2013 Update), and 3 years after the majority finished high school (second follow-up). HSLS:09 did not freshen the student sample; therefore, HSLS:09 is, for example, not representative of 11th-graders. Also, the representativeness of the school sample is lost after the base year.

HSLS:09 attempts to preserve the best design features of the predecessor high school longitudinal studies, while updating and improving upon those prior studies. The data collection points for HSLS:09 were chosen for their research value, considered independently of the data collection points employed in earlier secondary longitudinal studies. The base-year 9th-grade starting point was designed to capture—like NELS:88, which started in 8th grade—the transition into high school. It does so without the financial costs of following a sample in which 95 percent of the cohort had changed schools by the time of the first follow-up 2 years later, as experienced in NELS:88.

The HSLS:09 first follow-up took place when most students were in the spring term of 11th grade. It has often been observed that students in the spring of their senior year are disengaging from high school and not highly motivated to complete low-stakes assessments and questionnaires. Much thought has been given to improving students' participation and effort (e.g., as in NAEP, which traditionally has conducted 12th-grade as well as 4th- and 8th-grade assessments [see StandardsWork 2006]). One possible approach to addressing this problem is to move the testing point to spring of 11th grade, the strategy embraced by HSLS:09.

The timing of the 2013 Update—the last half of 2013 after (modal) graduation—also reflects a conscious choice. Earlier studies had data collections in the spring term—as early as January and February—of the senior year in high school, a time point at which many sample members had yet to make final decisions about postsecondary schooling or work. Much of the information about the decision process and its outcomes had to be collected, if at all, at the time of a follow-up 2 years after the senior year, when recollection of process details, including acceptances, rejections, and financial aid offers, had diminished. The Update's timing strengthens the HSLS:09 longitudinal design by collecting decision information immediately following typical graduation.

The timing of the second follow-up, with student survey administration beginning in 2016, likewise was based on specific research considerations. In the past studies, the interval between high school graduation and the follow-up questionnaire was 2 years. For HSLS:09, the interval was 3 years. One benefit of this longer interval was the opportunity to obtain better information on postsecondary education persistence and subbaccalaureate attainment. A second benefit was that, at the time of the second follow-up, the subsets of HSLS:09 and BPS:12/14 first follow-up students who were immediate postsecondary entrants were aligned in terms of the amount of time that had elapsed since beginning postsecondary education, allowing for comparisons to be drawn between these two cohort subsets. Both BPS:12/14 and HSLS:09 immediate postsecondary entrants were followed 3 years after first enrollment.

Although HSLS:09 offers the design benefit of important new measurement points, a trade-off should be noted. Specific cross-cohort comparisons cannot be made with the earlier secondary longitudinal studies. Nor can comparisons be made with the high school transcript studies of NAEP. HSLS:09 is based solely on a fall 9th-grade cohort, whereas the prior longitudinal studies were based on spring-term 8th-, 10th-, or 12th-grade cohorts (see figure 1). NAEP transcripts were collected only for graduating seniors and are nationally representative for that population. Similarly, the links between NAEP, NELS:88, and ELS:2002 mathematics assessments cannot be replicated within the HSLS:09 design.

A final point about the comparative structures of HSLS:09 and its two most recent predecessor studies pertains to sample "freshening," a device for cost-efficiently generating multiple grade-representative cohorts during a longitudinal study. As mentioned above, HSLS:09 includes only a single cohort, not two (grades 10 and 12 as in ELS:2002) or three (grades 8, 10, and 12 as in NELS:88); the 9th-grade student sample is the sole cohort across all rounds. The earlier studies freshened the sample to represent later grades. This was done for a compelling reason: to facilitate cross-cohort comparisons (e.g., trends among high school seniors in 1972, 1980, and 1992). Because HSLS:09 has no specific cross-cohort comparison points within the family of NCES secondary longitudinal studies, the traditional rationale for freshening does not apply. Freshening was also problematic because the 9th-grade sample does not represent all, or nearly all, 9th-graders—schools were eligible if and only if they had both a 9th grade and an 11th grade at the time of sampling.

This page intentionally left blank

Chapter 2. Sample Design

This section provides details of the sample design employed for the High School Longitudinal Study of 2009 (HSLS:09) second follow-up study. The second follow-up sample consisted of those study-eligible students selected for the base year in 2009–10 who were not deceased as of the second follow-up. Sections 2.1, 2.2, and 2.3 summarize the school and student sampling used for the base year, first follow-up, and 2013 Update and High School Transcript study, respectively. The student sample for the second follow-up study is described in section 2.4.

2.1 Base-year Sample Design

Selection of the school sample. HSLS:09 employed 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 in the second stage. The HSLS:09 target population of schools 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 that provided instruction to students in both the 9th and 11th grades as of fall 2009. For details of the rules for school inclusion or exclusion, see the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011). A total of 944 of 1,889 eligible schools participated in the base year, resulting in a 55.5 percent weighted school response rate (50.0 percent unweighted).

Although HSLS:09 was designed to be representative of 9th-grade students in the 2009–10 school year in study-eligible schools across the United States (i.e., a national design), it also supports construction of select state-level estimates for students enrolled in 9th grade in public schools in the fall of 2009. In particular, in response to 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: California, Florida, Georgia, Michigan, North Carolina, Ohio, Pennsylvania, Tennessee, Texas, and Washington. Additional information on construction of the HSLS:09 base-year school sample may be found in the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011).

Selection of the student and contextual samples. The student target population contained all 9th-grade students as of fall 2009 who attended either regular public or

private schools[8] in the 50 states and the District of Columbia that provided instruction in both 9th and 11th grades. This population is referred to as the "9th-grade cohort" in the subsequent discussions, where appropriate.

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. Student participants completed an in-school survey and mathematics assessment.

Contextual information was collected on the student sample to describe the home and school environments. Home life and background information was obtained through questionnaires completed by students' parents. Administrator and counselor questionnaires provided school information. 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.

For additional information on selection of the HSLS:09 base-year student and contextual samples, please refer to the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011).

2.2 First Follow-up Sample Design

The first follow-up student target population is the same as defined for the base year.

First follow-up student and contextual samples. All 25,206 base-year study-eligible students, regardless of their response and enrollment status, were included in the first follow-up sample. Unlike prior National Center for Education Statistics (NCES) high school longitudinal studies (NELS:88 and ELS:2002), the HSLS:09 student sample was not freshened to include a representative later-grade cohort, such as 11th-graders in HSLS:09. Therefore, first follow-up estimates from the sample are associated only with the 9th-grade cohort 2 1/2 years later, and not the universe of students attending the 11th grade in the spring of 2012.

[8] The term "regular" refers to the setting and mode of instruction. Some examples of schools not considered regular are those that offer instruction in juvenile detention centers, schools that instruct only special education students, and schools where all the students may be homeschooled or where a mix of instructional modes is used (e.g., some students are homeschooled, some receive remote instruction, and some are in a common physical location).

Some students were deceased as of the first follow-up, withdrew from HSLS:09 prior to the first follow-up, or were determined to be study ineligible for HSLS:09 as of the first follow-up. The number of sample members fielded and found to be eligible and the number found to be ineligible or deceased are provided in figure 4 below. Of the 25,206 eligible base-year sample members, 25,184 remained eligible as of the first follow-up.

The student questionnaire explored a variety of topics that include, but are not limited to, high school attendance, grade progression, school experiences, demographics and family background, completion of admission tests, college choice and characteristics, and high school coursetaking. Contextual information was collected for the student sample to describe their home and school environments. Home life and background information was obtained through questionnaires completed by students' parents. The first follow-up parent questionnaires were administered to the parents of a random subsample of students, whereas parent questionnaires were sought for all students in the base year. School information was obtained through the administrator and counselor questionnaires; however, administrator data were collected at both the base-year schools and the schools to which sample members transferred. Counselor data were collected in the first follow-up only from base-year high schools. For additional information on selection of the HSLS:09 first follow-up student and contextual samples, see the *HSLS:09 Base Year to First Follow-Up Data File Documentation* (Ingels et al. 2013).

2.3 2013 Update and High School Transcript Study Sample Design

In the 2013 Update, students or their parents responded to a survey in which information was collected on the student sample to describe the student's high school completion status, postsecondary education and work plans, college application experiences, and work experiences. In addition, school personnel in base-year schools and other schools identified during data collection supplied high school transcripts for HSLS:09 students from all schools that these students had attended.

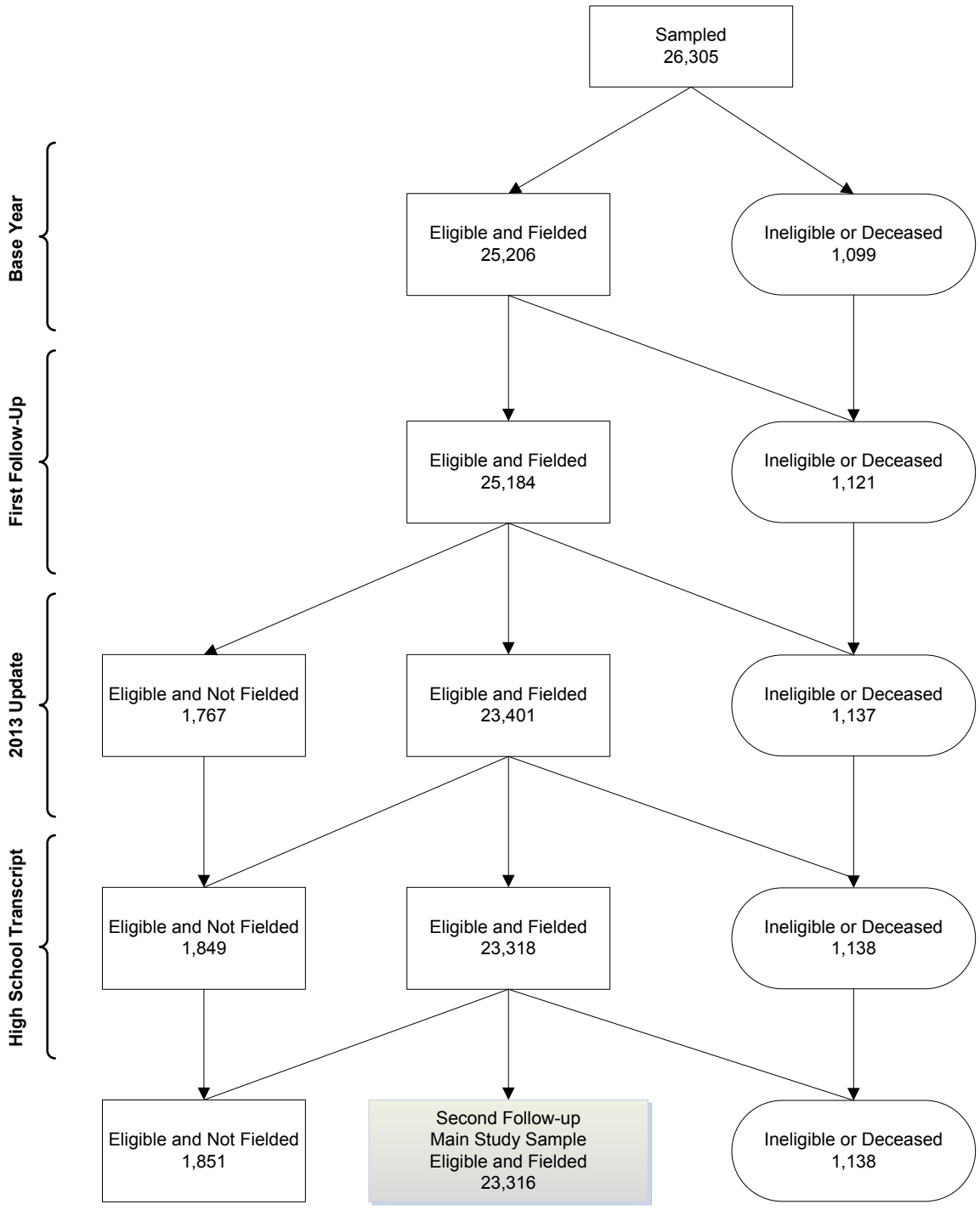
The number of sample members fielded and found to be eligible, not fielded but eligible, and found to be ineligible or deceased are provided in figure 4 below. Of the 25,206 eligible base-year sample members, 25,168 remained eligible as of the 2013 Update and 25,167 of the 25,206 eligible base-year sample members remained eligible as of the High School Transcript data collection.

Figure 4 shows that 1,767 of the first follow-up eligible and fielded sample members were not fielded for the 2013 Update. The majority of these sample members were nonrespondents in both the base year and the first follow-up. Additionally, some sample members were not fielded for the 2013 Update because they withdrew from the study. Information on selection of the HSLS:09 2013 Update sample appears in the *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015).

2.4 Second Follow-up Sample Design

The second follow-up fielded sampled included 23,316 of the 23,401 sample members fielded and found eligible for the 2013 Update. The 85 sample members not fielded withdrew from the study between the end of the 2013 Update collection and the beginning of the second follow-up data collection or were found to be deceased.

Figure 4. Student eligibility and fielding disposition from base year to the second follow-up: 2016



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

This page intentionally left blank

Chapter 3. Instrumentation

3.1 Base-year, First Follow-up, and 2013 Update Instruments

Base-year instruments. Conducted during the fall term when sample members were in the 9th grade, the base-year data collection consisted of a student survey and direct measure of math skills (i.e., algebraic reasoning), complemented by surveys of sample members' parents, math and science teachers, school counselors, and school administrators.

The student survey collected information on sample members' demographic characteristics and language use; school experiences from the fall 2009 term and previous academic year (e.g., mathematics and science experiences and course enrollment); mathematics and science self-efficacy; high school, postsecondary, and career plans; and other topics. For a full list of topics covered in the base-year student survey, see figure 3 in the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011). The direct measure of math skills was designed to focus on algebraic reasoning and was vertically scaled to show gains between the base year and first follow-up, when most sample members were in the 11th grade). For more detailed information about the direct measure of math skills, as well as details on the other survey components (i.e., sample members' parents, math and science teachers, school counselors, and school administrators) see section 2.2 and section 2.3 of the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011). Appendix A of the *HSLS:09 Base-Year Data File Documentation* provides specifications of each survey instrument.

First follow-up instruments. The first follow-up was conducted in the spring term of 2012 when most sample members were completing 11th grade, although sample members did not need to be on-grade or enrolled to remain part of the study. Each component from the base year was fielded again, except for the math and science teacher questionnaires. In all surveys, base-year themes were carried forward with new topics added, appropriate for rising seniors, early high school graduates, alternative completers (i.e., those who earned a GED test credential or a similar credential), and dropouts. For more detailed information about the surveys' contents and the direct measure of math skills during the first follow up, see sections 2.2 and section 2.3 of the *HSLS:09 Base Year to First Follow-Up Data File Documentation*

(Ingels et al. 2013). Appendix A of the *HSLS:09 Base Year to First Follow-Up Data File Documentation* provides specifications of each survey instrument.

2013 Update instrument. The goal of the 2013 Update was to collect information on sample members' status with respect to high school completion, postsecondary applications and enrollment, financial aid applications and offers, and employment. Two instrument design strategies were used to maximize the response rate. First, the instrument's response time was designed to average about 15 minutes, approximately half the length of the previous HSLS:09 instruments. Second, either the sample member or a parent could respond to the interview. Given this design, an effort was made to select objective questions that sample members and parents would respond to consistently. The data file includes a variable indicating whether the sample member or the parent was the respondent. Sample members' high school transcripts were also collected from base-year schools and schools to which sample members had transferred. For more detailed information about the 2013 Update survey or the High School Transcript data components, see the *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015). Appendix B of the *HSLS:09 2013 Update and High School Transcript Data File Documentation* provides specifications of the 2013 Update survey instrument.

3.2 Instrument Development in the Second Follow-up: Goals, Processes, Procedures

The goal for developing the content of the second follow-up instrument was to design a web survey that, in conjunction with data collected in previous rounds of the study, could provide insight into three areas of research interest for HSLS:09:

- What are students' trajectories from the beginning of high school into postsecondary education, the workforce, and beyond?
- What majors and careers do students decide to pursue when, why, and how?
- How do students choose STEM courses, majors, and careers?

A field test of the draft survey instrument was conducted prior to the main study. Development of the field-test instrument included two rounds of cognitive interviewing: one prior to deployment of the field test and a second round occurring in parallel with the field-test data collection. The cognitive interviews provided hands-on testing of the survey's programming functionality and usability using a variety of web-accessible platforms including mobile and nonmobile devices. Results of both rounds of cognitive testing are presented in appendix C.

Additional steps were taken during the field test to evaluate the survey. Analyses included evaluation of response distributions, item-level nonresponse, interview timing, and examination of test-retest reliability for about 30 selected items. Regular meetings were conducted with telephone interviewers during the field-test data collection to elicit their feedback on survey questions both in terms of ease of administration and respondents' ease of interpretation and response. Survey staff also monitored recordings of telephone interviews to identify questions that performed well and those that did not.

While the field-test data collection was underway, NCES project staff and the survey contractor began a series of instrument design meetings, with the goal of refining the field-test instrument to create the main study instrument. Based on the research goals of HSLS:09 and the results of the cognitive interviews and field test, preliminary recommendations for inclusion, exclusion, addition, and revision of items were developed and presented to a Technical Review Panel (TRP) composed of experts in secondary and postsecondary education. Feedback from the TRP was then taken under advisement as final decisions were made by NCES about the content and structure of the web survey. A subset of items considered to be most essential for the HSLS:09 research agenda were identified for inclusion in an abbreviated survey instrument to use at the end of data collection. Appendix B provides details on development, testing, and evaluation of the field-test instrument.

3.3 Survey Instrument Content in the Second Follow-up

The second follow-up data collection began about 3 years after the majority of the sample was expected to have graduated from high school. At the time of the second follow-up data collection, most respondents had completed high school having earned a high school diploma or another high school equivalency credential, though some had not. Respondents may have enrolled in a postsecondary institution, worked or served in the military, or formed a family. The second follow-up survey contained four substantive sections, grouped by broad topics:

- **High school.** Questions in this section primarily consisted of data elements included in previous study rounds for which data were missing for some respondents. Respondents could be missing these data due to nonresponse at either the unit level or item level, or because the question previously did not apply to them. To reduce sample members' time within the section, respondents for whom a data element was already available were routed around the corresponding question. Information that was collected to augment or update preexisting data included the type of high school

credential earned and the associated date, the last high school attended, high school grade point average, the grade in which algebra I was taken, the highest math course completed, whether the respondent had ever transferred high schools, and whether the respondent had ever dropped out of high school. Respondents who had not been awarded a high school credential were asked to report the grade they were in when they last attended high school, whether they had ever participated in a high school completion program, if they expected to earn a high school credential by the end of 2016, and if they had taken the test for the GED or another high-school equivalency credential. Those who had taken a test for a high school credential, including those who had earned a GED or other high school equivalency credential, were asked if they passed all components of that exam the first time they had taken the test. Recipients of GEDs or other high school equivalency credentials were also asked to report the state that awarded their credential. The last questions in the section asked whether the respondent had earned college credits while still enrolled in high school; if they had, the sample member was then asked to provide the name of the credit-granting postsecondary institution, so that transcript data could be collected.

- **Postsecondary education.** Questions from the 2013 Update that dealt with college applications were repeated and administered only to 2013 Update nonrespondents and respondents who had not applied to college at the time of the 2013 Update. Respondents were asked to identify the institution they eventually selected to attend, and up to two other institutions they most seriously considered attending. Respondents were also asked about their admission status at each institution reported, their first-choice institution, and the reasons they attended an institution other than their first choice, if applicable.

  All respondents who had enrolled at a postsecondary institution following high school, including those who had reported on college applications in the 2013 Update, were asked to provide a complete enrollment history through the second follow-up's reference date of February 2016. Information collected included the names of institutions attended, start- and end-dates for each institution, degrees or certificates pursued at each institution, completion dates or expected completion dates for each degree or certificate, and reasons for taking classes outside of a degree or certificate program, if applicable. Students were asked to report their enrollment intensity (i.e., full-time, part-time, or a combination of full-time and part-time) during their entire enrollment period. Sample members who indicated

delaying entry, dropping out, or transferring were also asked about their reasons for their enrollment pattern; similarly, respondents who reported not attending any postsecondary institution after high school were also asked their reason for not doing so.

All respondents, including those who had no postsecondary enrollment between high school and the second follow-up's reference period, February 2016, were asked to identify any previously unlisted institutions they had attended or planned to attend between February 2016 and the end of 2016. Postsecondary transcripts were collected from these institutions as part of the postsecondary transcripts data collection. Additionally, those respondents who had not enrolled in a bachelor's degree program by the end of February 2016 were asked if they planned to do so within the next 3 years.

Another series of questions asked of postsecondary enrollees related to majors or fields of study. These students were asked for their intended major or field of study, upon entry into their first postsecondary institution, and for the major or field of study for their "reference degree." The reference degree is identified on the data file through the variable X4REFDEG and was the undergraduate degree or certificate that the respondent was working on in February 2016, or most recently before that time. It could be a degree or certificate that the respondent was actively pursuing in February 2016, a completed postsecondary credential, or an incomplete degree or certificate that the respondent was no longer working on. Respondents who had declared or decided upon a major/field of study for their reference degree were asked to indicate what it was and why they had chosen the major or field of study. Those who had changed their major from what they had initially intended to study upon entry into postsecondary education were asked their reasons for doing so.

Respondents who had any postsecondary enrollment after high school were asked additional questions about their postsecondary experiences. These sample members were asked if they had taken STEM courses; how professors in their STEM courses treated male and female students, as well as students of different races; whether they had ever taken remedial courses or requested help for a course; whether they had used college services (e.g., financial aid counseling, academic support services); whether they

participated in high-impact activities[9] (e.g., study abroad, research project, or a community-based project as part of a course); whether they lived on campus; whether they had ever enrolled in a program that was entirely online; whether they had received private loans or scholarships; and the sample member's perceived value of their postsecondary education.

In addition to the questions described above, there were a number of questions from previous rounds of the study that were repeated for all respondents, including those who had not enrolled in a postsecondary institution after high school. All respondents were asked about their expected educational attainment. Additionally, given that a key area of focus in HSLS:09 is coursework and employment related to STEM, all respondents were asked about their ability in STEM.

- **Employment.** The third instrument section focused on two particular employment experiences: (1) the first job held after high school, and (2) the job held in February 2016 or most recently before February 2016. Information that was collected from all respondents concerning both jobs included job title and employer type (i.e., self-employed, military, or other), job start- and end-dates (month and year), whether the sample member was continuously employed between those dates, earnings, and hours worked per week while enrolled and while not enrolled.

  All respondents were also asked about the number of jobs they had held since high school, their participation in work experience programs (e.g., internships), whether they held a professional certification or a state or industry license, a detailed series of questions on military service (e.g., branches, active duty), and the value they place on various job characteristics relative to salary. Sample members were also asked to identify the job they expect to hold when they are 30 years old, how closely related the expected job would be to their job in February 2016 or most recent job, and their expected earnings at the future job. All respondents were also asked if they felt that they had been treated unfairly in education or in the workforce.

  Respondents who had postsecondary enrollment were asked additional questions about their employment while enrolled, including the number of hours they worked while enrolled, by academic year; interference of this

[9] Certain activities are designated as high-impact activities or practices due to their positive relationship with student learning and student retention. High-impact activities often demand substantial time and effort; facilitate learning outside the classroom; and encourage meaningful collaboration with faculty, students, and others (National Survey of Student Engagement 2017).

work on academic performance; and whether they identified primarily as a student or as an employee.

Respondents not enrolled in an undergraduate degree or certificate program in February 2016 were asked additional questions about the job they held in February 2016 or their most recent job. These respondents were asked to report the benefits offered by their employer, their overall job satisfaction, and whether they would have preferred to work more hours, regardless of full-time or part-time status. These sample members were also asked if the job was an apprenticeship or if it required licensure. Additionally, these respondents were asked to report the number and duration of any unemployment spells and their receipt of unemployment compensation.

- **Family and community.** This section asked all respondents about a range of topics including family and home life (e.g., marital and parental status, household composition), financial well-being (e.g., expenses, number of dependents, financial assistance to and from others, income, spouse's income, worries and behaviors related to money), community engagement (e.g., voting registration, volunteering), personal characteristics (e.g., disabilities, birth sex, gender identity, and sexual orientation), and significant life events (e.g., job loss, death of friends or close relatives, serious injury or illness). Foreign-born respondents who were not known to be U.S. citizens prior to the second follow-up were also asked if they were U.S. citizens in February 2016.

In addition to the substantive sections outlined above, the survey included a section at the end to collect locating information for any potential subsequent follow-ups with this cohort.

For further detail on the topics covered by the survey, please refer to the instrument survey specifications in appendix D. The survey specifications indicate to whom the question applied, as well as provide question wording, response options, and clarifying notes for analysts.

3.4 Survey Instrument Design and Features in the Second Follow-up

Respondents were asked to report on their activities through the end of February 2016, the month before data collection began, and many questions asked specifically about the respondent's activities and status in that month. By establishing

February 2016 as the survey's reference date, all respondents reported on a uniform time period regardless of when they completed the survey during the 11-month data collection window. If the survey had asked about current activities, two respondents who began postsecondary education or employment during the data collection period would have provided different responses if one completed the survey before the activities began and the other completed the survey after. The February 2016 reference date ensured comparability of the data across all respondents. See the variable X4SQDATE for the date the survey was completed.

Modes of administration. The survey was developed as a web instrument that could be used for self-administration as well as interviewer administration via computer-assisted telephone interviewing (CATI), in which the interview was administered over the telephone by an interviewer, or computer-assisted personal interviewing (CAPI), in which a field interviewer (FI) conducted the interview in person or by telephone. The web instrument could be completed on a desktop/laptop computer or a mobile device such as a tablet or smartphone. The survey did not have to be completed in a single session or in a single mode. Respondents could start the survey, log out, and resume where they left off. Moreover, respondents could switch modes across survey sessions. For example, they could begin the survey on their own and later complete the survey with an interviewer.

When the survey was accessed from a mobile device, the screen automatically adjusted for optimal viewing on a small screen. The survey banner was reduced in size, and nonessential items (e.g., the survey progress bar) were removed from the main screen, although they remained accessible through a menu. Moreover, questions that had a grid layout were reformatted to reduce the likelihood that they would extend beyond the screen borders. Particular attention was given to testing the survey on a variety of smartphones.

Minimizing potential mode effects. The survey instrument incorporated several design features to provide self-administered respondents the same assistance provided by an interviewer to minimize the potential that the mode of survey administration would influence survey responses. These included:

- help text to define key terms and clarify question intent;
- pop-up messages to correct responses that were out of range or incorrectly formatted;
- conversion text to encourage responses to critical items when these items were left unanswered; and

- pop-up messages prompting sample members to provide a response when three consecutive questions were left unanswered.

For sample members who completed the survey over the phone, trained interviewers administered the survey. For computer-assisted telephone interviews and computer-assisted in-person interviews, the survey instrument included instructions for interviewers on each screen indicating how each question was to be administered, such as whether the response options were to be read aloud and when to probe for more information.

Survey length. In developing the survey instrument, a primary challenge was to achieve the desired length for the interview while retaining as many data elements as possible. This was achieved in large part by differentially routing respondents based on their enrollment status and employment status. For example, respondents who were enrolled in a postsecondary institution, but had not or were not currently working, had a much shorter Employment section compared to their peers who had worked. Survey length was also reduced by routing respondents around questions when the answer could be logically inferred from the answer provided to an earlier question. For example, if a spring 2013 high school graduate indicated that she or he was working for pay while enrolled in a postsecondary institution during the 2013–14 school year (i.e., S4WRK1314 = 1), then the subsequent question asking if the student had worked for pay since high school (i.e., S4ANYJOB) did not need to be asked. For further information on logical inference, see section 5.2.

Use of full-length and abbreviated surveys. The full-length survey instrument was used exclusively from the beginning of data collection through December 11, 2016, at which time it was replaced by the abbreviated instrument to encourage responses from the remaining nonrespondents. Survey respondents who partially completed the full-length instrument before that date and resumed after that date were automatically switched from the full-length to the abbreviated survey. The variable X4SQSTAT indicates whether the survey was a full-length survey, an abbreviated survey, or a combination of both. For a distribution of interview responses by data collection phase, see table 11 in section 4.3.2; for an indication of which items were included in the abbreviated interview, see the survey specifications provided in appendix D.

3.5 Criteria for Defining Completed Interviews in the Second Follow-up

Due to their analytic importance, certain items factored heavily in response status determination. Because of the nature of the survey, in which respondents had the ability to answer or skip any item and could potentially respond to an abbreviated version of the instrument, completeness of data varies across respondents. The amount of information required for inclusion on the data file reflected a dual requirement which depended on a combination of the amount of data provided and the substantive value of the data provided. A case was defined as a study respondent if the sample member answered a combination of a certain percentage of critical items and provided a core amount of enrollment or employment data.

Specifically, a sample member was considered a respondent if either of the two following criteria were met:

1. The respondent had data for all critical items (i.e., X4HSCOMPSTAT [High school credential status and type, February 2016], X4HSCOMPDATE [Date received high school credential], S4EVRATNDCLG [Ever attended college by the end of February 2016], and S4WORKING16FB [Working for pay in any job in February 2016]).
2. The respondent had data for at least one of the critical items and
    a. identified a postsecondary institution that was not identified in the 2013 Update; or
    b. identified a postsecondary institution that was identified in the 2013 Update and
        i. reported completing a postsecondary credential; or
        ii. specified a major field of study; or
        iii. indicated when last attended a postsecondary institution; or
    c. did not identify a postsecondary institution but did indicate work status; or
    d. indicated not attending a postsecondary institution and provided new high school credential information.

Chapter 4. Data Collection Methodology and Results

Chapter 4 describes data collection design, procedures, and results for the High School Longitudinal Study of 2009 (HSLS:09) second follow-up. Data file documentation reports for the base year, first follow-up, and 2013 Update and High School Transcript study provide detailed descriptions of the procedures and results for the base-year, first follow-up, and 2013 Update and High School Transcript data collections conducted in 2009, 2012, and 2013, respectively (see Ingels et al. 2011, 2013, 2015).

4.1 Data Collection Methodology

HSLS:09 second follow-up data collection was conducted from March 2016 through January 2017 and consisted of administering a full-length interview averaging 32 minutes and a 17-minute abbreviated interview through several modes: self-administered via the Web, computer-administered telephone interviewing (CATI), and computer-assisted field interviewing (CAPI).

This section summarizes data collection procedures implemented in the second follow-up. The section describes resources available to sample members, including the study website and help desk. It also describes data collection staff training; tracing, locating, and interviewing procedures; and quality control procedures.

4.1.1 Website

The HSLS:09 website provided general information about the study, including answers to frequently asked questions (FAQs), information about confidentiality and data security, and selected analytical findings from previous rounds of data collection. The website also contained contact information for the study help desk and project staff at RTI, as well as links to the main National Center for Education Statistics (NCES) and RTI websites. A link to the study website was provided in all mail and e-mail communications with sample members.

The study website also hosted the self-administered web interview, which was available continuously throughout the entire data collection period. Data collection e-mails and mailings encouraged sample members to complete the web interview,

although the telephone interview remained available for sample members who called the help desk. The web and telephone interviews were identical, apart from on-screen instructions that were provided to telephone interviewers.

The study website included several features designed to protect sample members' data. All sample members were provided with secure log-in credentials composed of a unique study ID and strong password with which to access the interview. Any data provided on the website were protected using a Secure Sockets Layer (SSL) protocol that allowed only encrypted data to be transmitted over the Internet. All data collected were stored in a secured SQL server database located on an NCES server that was physically separate from the web server. Figure 5 shows the log-in page for the HSLS:09 second follow-up website.

Figure 5. HSLS:09 second follow-up website: 2016

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

4.1.2 Help Desk

HSLS:09 staffed a help desk with data collection interviewers trained to assist sample members accessing the web interview, answer questions about the study, and provide sample members with their log-in information. Study members could reach the help desk using either a toll-free telephone number or e-mail. The toll-free telephone number and e-mail address were provided in all materials sent to sample members.

Sample members primarily contacted the help desk to request their study ID, a new password, or to complete the interview by telephone. For each inbound call to the help desk, staff confirmed the identity of the caller and logged the results of the call in the computer-assisted telephone interviewing case management system (CATI-CMS).[10] The CATI-CMS enabled interviewers to send an e-mail reminder, containing the sample member's study ID and password, to those who wished to complete the self-administered web interview at a later time. Additionally, for technical issues that could not be immediately resolved, help desk staff offered to administer the interview over the phone.

4.1.3 Data Collection Staff Training

Members of the second follow-up data collection team filled several specified roles, and all completed comprehensive study-specific and role-specific training programs before beginning work on the study. Training sessions included instruction on the background and purpose of HSLS:09, confidentiality procedures, case management procedures, and FAQs, as well as hands-on activities designed to maximize familiarity with and proficiency in their respective tasks.

The training schedule and number of data collection staff members trained for each role are presented in table 1.

[10] See section 5.1 for a description of the computer-assisted telephone interviewing case management system (CATI-CMS).

Table 1. Data collection staff trainings: 2016

Staff trained	Time period	Number of staff trained
Data collection interviewers, performance team leaders, and quality experts	March 15–17, 2016; May 17–19, 2016; and September 13–15, 2016	56
Field interviewers and field supervisors	August 30–September 1, 2016 and November 17–19, 2016	44
Tracing staff	May 9, 2016; August 3, 2016; and January 10, 2017	30

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

Data collection interviewers (DCIs). DCIs were the primary point of contact with sample members, and their responsibilities included conducting telephone surveys, handling incoming help desk calls, addressing sample members' concerns, helping sample members log in to the web survey, gaining cooperation, and converting refusals.

Prior to the start of data collection, all DCIs received 12 hours of in-person instruction that included general interviewing procedures as well as study-specific training. General instruction covered professional interviewing techniques, confidentiality procedures, sample member rights, and call center procedures. Following general instruction, study-specific training included an overview of HSLS:09; a review of the second follow-up survey instrument through hands-on practice with mock interviews and the CATI-CMS; and guidance on providing technical support to sample members. Data collection managers certified DCIs to begin work on the study after the interviewer successfully conducted a mock interview and provided appropriate and accurate responses to FAQs.

Performance team leaders (PTLs). PTLs supervised and oversaw DCI performance, provided guidance to interviewers, and helped troubleshoot escalated problems. PTL training included the content covered in the interviewer training and additional training in case review, problem resolution, project-specific reporting, and other procedures specific to HSLS:09.

Quality experts (QEs). QEs monitored live and recorded interviews and provided feedback and coaching to interviewers on their performance. QEs attended interviewer training to learn the content of the second follow-up survey, standards for administering the survey, and project-specific procedures. They also received additional training on how to evaluate interviewers' performance and provide feedback.

Field interviewers (FIs). Before attending in-person training, FIs were asked to review the HSLS:09 field interviewer manual, review an online training module that provided background information on HSLS:09 and study protocols, and complete an in-home exercise. The 3-day in-person training included topics such as locating sample members in the field, gaining cooperation, using the Integrated Field Management System (IFMS)[11], administering the interview, study FAQs, and use and care of the field laptop computer (laptop). Before commencing work on the study, each FI was required to pass a series of certification assessments to demonstrate mastery in each of several skill areas.

Field supervisors (FSs). FSs completed all in-home and in-person training provided to FIs and received additional training on field tracing, study protocols, managing caseload, and handling difficult situations.

Tracing staff. Tracing staff received between 4 and 8 hours of training on general tracing procedures, depending on each staff member's level of experience. Tracing staff then received 2 more hours of HSLS:09-specific training, including an overview of HSLS:09, the second follow-up, FAQs, and intensive tracing techniques most appropriate for locating HSLS:09 sample members.

4.1.4 Tracing, Locating, and Interviewing Procedures

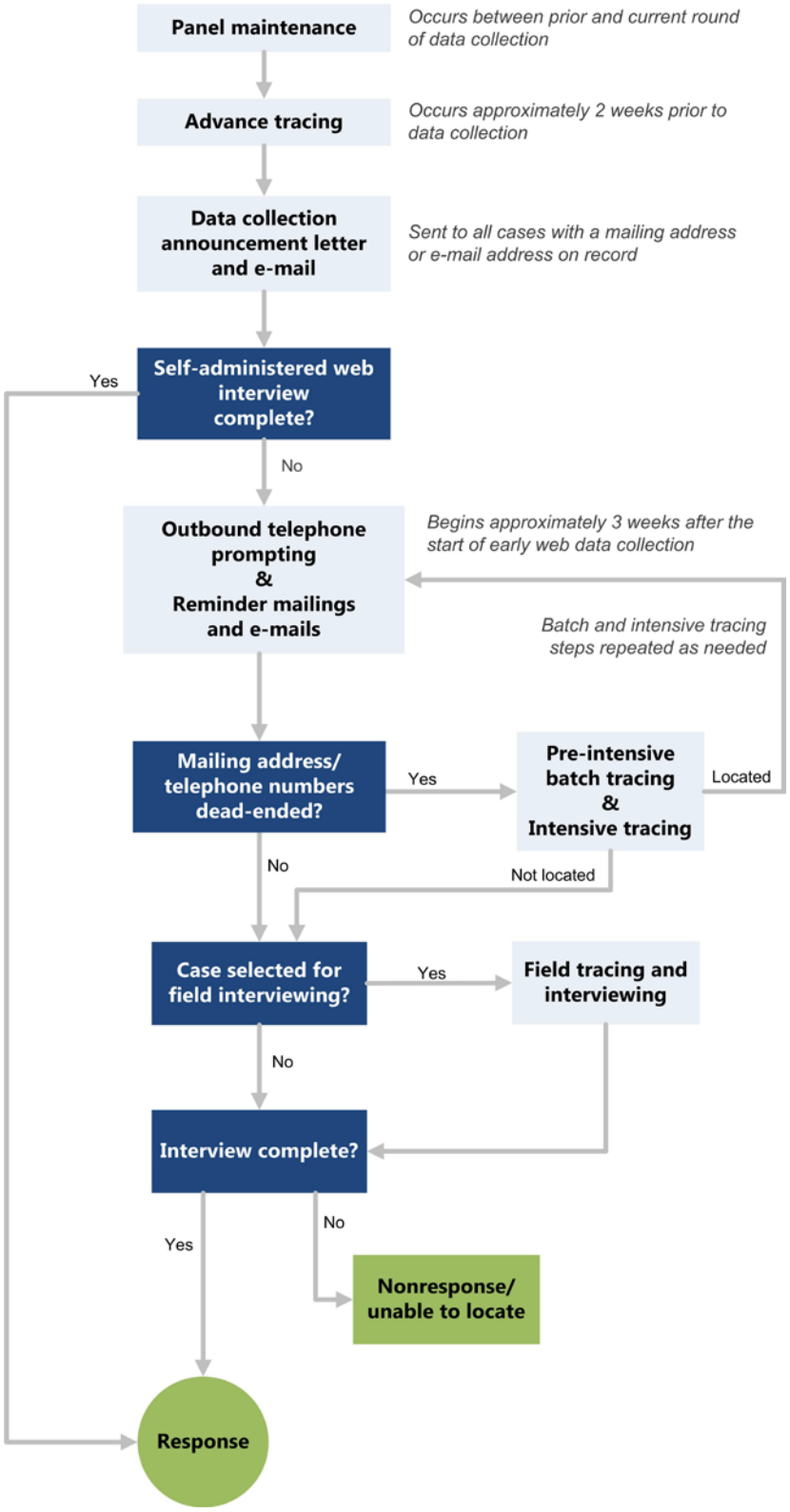
The HSLS:09 second follow-up used a multistep process for locating, tracing, and contacting sample members. These steps, described in detail below, included a panel maintenance activity to update sample members' contact information before the start of the second follow-up data collection. Then, closer to the start of data collection, project staff used several batch-tracing databases to update or confirm sample members' contact information. Sample members who did not have sufficient contact information on file after this initial batch tracing underwent additional batch tracing and intensive tracing effort prior to the start of data collection. Intensive tracing continued throughout data collection on a flow basis, for sample members whose contacting leads were exhausted. For sample members who were selected for field interviewing as part of data collection phase 5,[12] field staff conducted additional tracing by using local resources to develop contacting leads.

Once contact information was obtained for sample members, they were contacted by mail and e-mail and encouraged to complete the interview. Sample members had the option to complete a self-administered web interview, which was optimized for mobile devices, or a telephone interview. Prompting calls were made to sample

[11] See section 5.1 for a description of the Integrated Field Management System (IFMS).
[12] For descriptions of each data collection phase, see section 4.2.1.4.

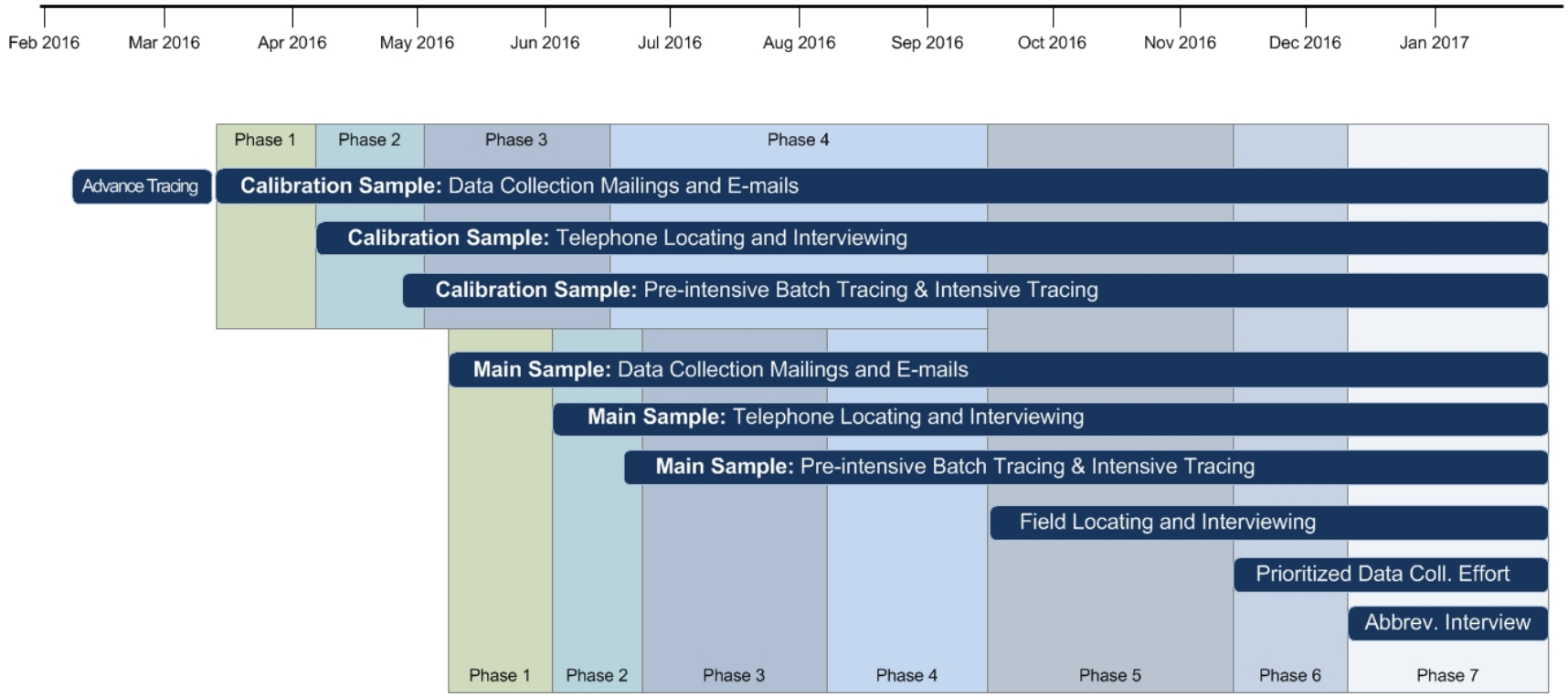
members who did not respond to the initial mail and e-mail requests. Mail, e-mail, and telephone follow-up continued throughout data collection. Sample members selected for field interviewing were also contacted in person and had the option to complete an in-person interview. Figure 6 illustrates the steps used to locate, trace, and contact sample members. Figure 7 illustrates the timeline of data collection phases and activities. Further details of the data collection phases as well as the calibration sample and main sample are provided in section 4.2.

Figure 6. Tracing and locating procedures: 2016



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

Figure 7. Timeline of data collection phases and activities: 2016



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

4.1.4.1 Panel maintenance

Beginning in August 2015, 7 to 9 months before data collection, a panel maintenance activity was conducted to confirm and update sample members' contact information. First, batch tracing was conducted using several sources to update contact information for sample members and one primary parent.[13] After the batch tracing efforts, a panel maintenance mailing was sent to sample members and parents, asking them to provide updated contact information and including a $10 incentive offer, payable to the sample member. Sample members or their parents could update contact information by completing a form on the HSLS:09 website or by calling the toll-free telephone line; parents were also sent a paper contact update form that they could fill out and return in a business-reply envelope provided.

After the panel maintenance mailings were sent, cases with the least complete contact information were identified for outbound telephone prompting. Approximately 3,500 cases received telephone contact efforts. After several weeks of telephone prompting, intensive tracing was conducted on a subset of cases with insufficient locating information to receive data collection announcement mailings. Approximately 3,000 cases received intensive tracing during the panel maintenance period. At the end of panel maintenance in January 2016, confirmed or updated contact information had been obtained for 33 percent of the cases.

4.1.4.2 Advance tracing

The advance tracing stage occurred approximately 2 weeks before the start of data collection and primarily included batch database searches. All cases with information sufficient for matching were included in batch tracing, regardless of whether they already had contact information on file. For cases with contact information already on file, batch tracing provided new or updated information and helped to identify the most current information. For the subset of cases that still did not have a valid mailing address after batch tracing, advance tracing also included intensive tracing, in which cases were sent to additional databases and tracing staff searched for contacting leads, as described in section 4.1.4.

Some searches were conducted for all sample members, regardless of whether they had contact information on file, and other higher-cost searches were only conducted

[13] In the second follow-up, parents were contacted as part of sample member locating efforts but were not asked to participate in a survey.

for sample members whose information did not match to previous sources. General descriptions of the search databases follow.

LexisNexis National Change of Address (NCOA) Database. NCOA is a database consisting of change-of-address data submitted to the U.S. Postal Service (USPS). Almost 160 million records are updated weekly and stored for 4 years. Searches may result in new addresses or confirmation of existing address records.

LexisNexis Phone Append. Phone Append is a computerized lookup service consisting of over 170 million residential telephone number listings, over 6 million not-yet-published numbers of new movers, and over 14 million businesses. Phone Append provides both cellular and landline phone numbers, and uses names, street addresses, and ZIP codes as search criteria.

LexisNexis Single Best Address Database. Although NCOA provides information only for people who registered a change of address with USPS, Single Best Address can provide new addresses, including those not registered with NCOA. Single Best Address searches multiple data sources using progressive search logic to return the most current address available.

LexisNexis E-mail Search. The E-mail Search provided by LexisNexis is currently the only known batch product that provides e-mail addresses. The search uses a name, best-known mailing address, best known phone number, and Social Security number (SSN; not required) to search for e-mail addresses associated with the input information.

LexisNexis SSN Search. The SSN Search looks through multiple data sources using progressive search logic to return the most current phone numbers available. The search requires an SSN and uses this in conjunction with a name and best-known mailing address.

Experian TrueTrace. Experian TrueTrace uses Experian's credit database to provide the most current phone numbers and addresses available. Experian TrueTrace uses name, best known mailing address, SSN, and all known phone numbers to confirm or provide the best-known information. Unlike the SSN Search, Experian TrueTrace uses the input phone numbers and provides a confirmation of the phone numbers that match their database and provides phone numbers in their database that were not provided in the input file. The search also confirms whether the input mailing address matches their database and provides up to five additional mailing addresses.

First Data Premium Phone. Premium Phone searches over 475 million landline, Voice over Internet Protocol (VoIP), and wireless numbers in the United States, Puerto Rico, and Canada. Premium Phone was used as an intermediate step before intensive tracing.

National Student Loan Data System (NSLDS). NSLDS is the U.S. Department of Education's central database for federal student aid. NSLDS matching can provide name, address, telephone, and e-mail address information. HSLS:09 staff used NSLDS matching services to collect sample member locating data partway through the data collection period.

4.1.4.3 Intensive tracing

Sample members that could not be located by other methods were pursued through intensive tracing. These included cases that had no phone number to load into the CATI-CMS or for whom all known phone numbers were exhausted during outbound calling. Intensive tracing consisted of a multitiered approach in which the most cost-effective steps were taken first.

Prior to intensive tracing, cases were initiated for several batch tracing searches, including First Data Premium Phone, Experian TrueTrace, and LexisNexis SSN Search, depending on the availability of information necessary for matching. If new contact information was found through batch tracing, the case was returned to outbound dialing without continuing to intensive tracing.

The first tier of intensive tracing operations identified sample members in consumer credit bureau databases. If this search resulted in a new telephone lead, the case was returned to outbound dialing. If the search resulted in a new address only, tracing staff used directory assistance searches to locate an associated telephone number.

If not located in the first tier, cases were moved to the more intensive second tier of tracing operations. In the second tier, tracing staff conducted a thorough review of each case and determined the appropriate next steps based on the leads developed from prior tracing and contacting activities. Tracing staff searched consumer databases and other sources on a case-by-case basis for potential leads to reach the sample member or other contacts. Tracing staff finalized cases as unlocatable only after exhausting all leads.

4.1.4.4 Data collection mailings and e-mails

Approximately 3 weeks before the start of data collection, sample members and parents were sent a pre-data-collection contact mailing and e-mail informing them of

the upcoming start of data collection. The pre-data collection contact mailing was sent to the calibration sample[14] in February 2016 and to the main sample in April 2016.

Data collection announcement mailings and e-mails marked the beginning of data collection. The data collection announcement mailing included a letter that notified sample members of their selection for HSLS:09, requested that they complete the web interview, provided instructions and log-in information for accessing the web interview, and for those sample members eligible for an incentive, included information about the baseline incentive to be offered for completing the interview. The mailing also included a study brochure that explained the background and purpose of HSLS:09, information about confidentiality and data security, and contact information for the study help desk. The data collection announcement e-mail included the same content as the letter, with the addition of a direct link for accessing the interview. The data collection announcement mailing and e-mail were sent to the calibration sample in March 2016 and the main sample in May 2016.

Frequent reminder e-mails, postcards, and flyers were sent throughout the remainder of data collection. These materials were used to inform sample members of increased incentive offers at the start of data collection phases 3 and 4, and of the abbreviated interview offer. Table 2 lists the hardcopy mailings sent to sample members or their parents. Appendix E includes examples of contact materials sent to sample members throughout data collection.

[14] See section 4.2 for a description of the calibration and main samples and incentive structures.

Table 2. Hardcopy mailings: 2016

Mailing	Number of cases sent mailing	Percent of total fielded sample sent mailing[1]
Pre-data-collection contact letter—sample members	13,754	59.0
Pre-data-collection contact letter—parents	14,495	62.2
Data collection announcement letter[2]	22,047	94.6
Reminder postcard 1	15,595	66.9
Reminder postcard 2	14,342	61.5
Reminder postcard 3	11,927	51.2
Reminder postcard 4	10,775	46.2
Reminder postcard 5	9,049	38.8
Reminder postcard 6	7,055	30.3
Reminder flyer	6,041	25.9
Reminder postcard 7	5,486	23.5
Parent reminder	4,811	20.6
Reminder postcard 8	4,567	19.6
Final flyer	4,368	18.7

[1] Total fielded sample = 23,316.
[2] Note that because new addresses were acquired or updated through tracing over the course of data collection, the number of cases mailed the data collection announcement letter does not represent the total number of cases located during data collection.
NOTE: Includes hardcopy mailings sent to sample members and their parents or guardians. Mailings are listed in the order in which they were sent. Frequent reminder e-mails were also sent throughout data collection to sample members with an e-mail address on record.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

4.1.4.5 CATI locating and interviewing

Data collection for the second follow-up consisted of an early response period and an active contacting period. During the early response period, corresponding to data collection phase 1, sample members were encouraged to complete the self-administered web interview, and no outbound telephone prompting occurred, although interviewers staffed the help desk and were available to receive inbound calls. For most sample members, the early response period lasted 3 weeks.[15] During the active contacting period, DCIs actively called sample members to encourage them to complete the survey online or over the phone. In addition to conducting interviews, interviewers also attempted to locate hard-to-reach sample members and obtain contact information from parents or other contacts.

Contact attempts were managed using the CATI-CMS. Prior to the start of data collection, all available telephone contact information for sample members, parents,

[15] The early response phase lasted 1 week for high school noncompleters (i.e., cases within subgroup A) to move these cases more rapidly to the active contacting period with active prompting. Details describing the composition of subgroup A cases are provided in section 4.2.1.1.

and other contacts was loaded into the CATI-CMS. Phone numbers were prioritized such that the most recent information was attempted first. A call scheduler program weighed several factors to determine the most appropriate time at which to place the next call for each case and which number to use. Factors included any scheduled appointments, the timing and results of previous call attempts, and the number of calls made to each telephone number. The CATI-CMS maintained a comprehensive record of all contact attempts, the results of each attempt, status codes, and interviewer case notes, which interviewers updated after each contact attempt. If the sample member was reached and agreed to participate, interviewers launched the web interview directly from within the CATI-CMS.

4.1.4.6 *Field locating and interviewing*

Field interviewing began in September 2016—data collection phase 5—for remaining nonrespondent cases in the full sample (i.e., the combined calibration and main samples) that were targeted as part of the responsive design; see section 4.2.1.4 for details on the selection process. FIs were recruited from geographic areas with the highest current concentrations of HSLS:09 sample members. FIs were located across the United States, with several traveling FIs designated to visit areas that had significant numbers of sample members but not enough to justify hiring a local interviewer.

FIs managed their caseload using RTI-issued laptops, which were equipped with the IFMS. The IFMS was loaded with a full record of the case history, including all known contact information for the sample member and other contacts, information about participation in prior rounds of the study, records of all previous contact attempts by DCIs, and a record of any tracing conducted on the case. FIs reviewed the case information and drew upon their experience and training to develop leads for locating sample members, gain cooperation, and complete the interview. To minimize costs associated with expensive in-person contact attempts, FIs attempted all known phone numbers for a case before working in the field.

The field laptops launched the same interview used by DCIs, except that the interview was hosted locally on the laptop instead of on the study website. This allowed FIs to complete the interview in any location without the need for internet access. The interview could be completed with an FI in person or over the phone. Sample members could also ask to complete the self-administered interview on the FIs' laptops. Each day that interviewers worked, they synchronized their laptops with RTI's data collection systems such that completed cases were transmitted back to RTI and any new contact information was transmitted to the field laptops. Data stored on field laptops were encrypted and password protected.

4.1.5 Data Collection Quality Control Procedures

Quality control procedures included live, real-time monitoring of DCIs; monitoring of recorded interviews; field interview verification; quality circle (QC) meetings with data collection staff; and debriefing meetings to identify potential areas for improvement of data collection protocols and systems.

4.1.5.1 Interviewer monitoring and verification

Project and call-center staff monitored and evaluated DCIs' performance throughout data collection. Interviewer performance was monitored using a quality management system that allows for monitoring interviewers during the call or after the call using a review of audio recordings. The quality monitoring system provided tools for selecting interviewers, observing their work, evaluating performance, providing feedback, and analyzing performance data across interviewers. Interviewers were evaluated for professionalism, proper question administration, and clear pace and enunciation of speech. Approximately 7 percent of all telephone interviews across all shifts were evaluated.

FIs were evaluated for professionalism and to detect any signs of interview falsification. Project data collection managers and FSs monitored interviewer production to ensure that no interviewers had unusually high response rates or other suspicious patterns of response. Data collection managers and FSs also placed verification calls to a subset of sample members who completed the interview with a FI. In the verification call, RTI staff confirmed the date and time the interview was completed; that the interview was completed with the sample member and not another person in the household; that the sample member received his or her incentive, if applicable; and that the interviewer completed the interview in a professional manner. Verification calls were conducted for 10 percent of cases completed by each FI.

4.1.5.2 Quality Circle meetings

Project data collection managers conducted regular QC meetings with both telephone interviewing and field interviewing staff. QC meetings provided an opportunity for managers to share updates about data collection progress and for DCIs and FIs to share their experiences and ask questions of the project managers.

QC meetings with DCIs were used to provide additional training on a variety of topics, such as building skills for administering questions identified during monitoring, reinforcing successful techniques for gaining cooperation, and providing corrective feedback. The meetings often included team-building exercises and

hands-on practice activities. QC meeting facilitators prepared notes to summarize the topics discussed at each meeting, and interviewers were responsible for reviewing all notes. The notes served as a reference and resource for interviewers throughout the course of data collection.

Weekly meetings with FIs were used to discuss strategies for gaining cooperation from sample members and from contacts who could provide leads for reaching the sample member (*gatekeepers*), plans for coordinating interviewer travel and managing workload across interviewers, and proper techniques for administering interview questions. The FIs hired for the second follow-up had a wide variety of prior experience working as FIs. The meetings allowed them to discuss challenges encountered in the field, brainstorm strategies for overcoming them, and learn from other interviewers' expertise.

4.1.5.3 Interviewer debriefings

At the end of data collection, data collection managers conducted debriefings with DCIs, PTLs, QEs, FIs, and FSs to learn more about their experiences working on the study. Comments and discussion from staff identified areas of success during training and data collection and identified areas for improvement in future studies. Staff provided feedback on several broad areas, including the following:

- DCI and FI training;
- case management systems;
- strategies for dealing with gatekeepers and gaining sample member cooperation;
- survey administration; and
- QC meetings.

Data collection staff had positive feedback about their experience working on the second follow-up and provided suggestions for how future studies could be improved.

4.2 Responsive Design Methodology

A key data collection strategy employed during the second follow-up was the use of a responsive data collection design, or *responsive design*. Responsive design approaches seek to: (1) pre-identify key design features of a survey that could affect cost or data quality; (2) monitor those features throughout data collection; and (3) alter features of the design based on decision rules (Groves and Heeringa 2006). In previous rounds of HSLS:09 and in other NCES studies (e.g., BPS:12/14, B&B:08/12, and

the ELS:2002 third follow-up), responsive design methods have been used to improve key features of data quality and reduce costs.

The HSLS:09 second follow-up data collection sought to maximize data quality through a data collection approach designed to reduce variance between the responding sample and the overall sample, thus improving overall sample representativeness. The approach allowed the project team to determine, during data collection, how representative the responding sample was of the total sample, so that resources could be focused on gaining cooperation from the cases most needed to achieve balance in the responding sample. Plans for the HSLS:09 second follow-up were informed by results of prior data collection rounds and results of incentive experiments and responsive design modeling simulations from the HSLS:09 second follow-up field test. (Appendix B provides a description of the field test responsive design methods and experiments conducted to evaluate the effectiveness of various interventions.) Additionally, the design was informed by results of related longitudinal studies such as BPS:12/14, B&B:08/12, and ELS:2002 third follow-up.

To encourage sample member participation, various interventions—both monetary and nonmonetary—were deployed as part of the responsive design strategy. Monetary interventions included an initial *baseline* incentive offer and two incentive increases, or incentive *boosts,* offered to nonrespondents. Nonmonetary interventions included outbound CATI prompting, field interviewing, prioritized data collection, and an abbreviated interview which decreased the time-commitment required to complete the survey.

A distinguishing aspect of the HSLS:09 second follow-up responsive design approach was the use of specifically identified subgroups of interest, based on prior experience with the cohort: high-school noncompleters (subgroup A), "ultra-cooperative" respondents (subgroup B), and all other cases (subgroup C). The subgroups were created so that customized interventions could be tested and applied to each group independently. Two models were used to help identify, or target, cases for specific interventions. The models consisted of a *response likelihood* model to assign an estimated a priori probability of response for each member and a *bias likelihood* model to identify nonrespondents in underrepresented groups. To allow interventions to be further tailored, the data collection period was separated into seven distinct phases in which corresponding monetary and nonmonetary interventions could be applied to remaining nonresponding sample members. To evaluate the effectiveness of interventions used in each phase, the second follow-up data collection included a *calibration* sample in which a subset of the sample was used to identify the optimal intervention to be implemented in the *main* sample, for each phase. About 14 percent of the HSLS:09 second follow-up sample members, or

3,300 cases, were randomly selected to participate in experiments as part of the calibration sample, with data collection beginning in mid-March 2016. The calibration sample began data collection approximately 8 weeks before the main sample, and after each intervention with the calibration sample, the resulting response rates were examined and used to identify the most effective interventions for the main sample.

This section describes the design that was implemented for the second follow-up data collection and describes the models used to identify cases for targeting. Additional details on the development of the response likelihood and bias likelihood models, the effectiveness of each model, the results of the calibration sample experiments, and the effects of the responsive design approach on key survey estimates may be found in appendix F.

Table 3 lists the major dates and milestones for the calibration sample and main sample during data collection.

Table 3. Data collection schedule: 2016

Phase	Calibration sample	Main sample
Phase 1 (baseline incentive)	March 14, 2016	May 9, 2016
Phase 2 (outbound CATI)	March 21, 2016 (subgroup A) and April 4, 2016 (subgroups B and C)	May 16, 2016 (subgroup A) and May 31, 2016 (subgroups B and C)
Phase 3 (incentive boost 1)	May 4, 2016	June 20, 2016
Phase 4 (incentive boost 2)	June 15, 2016	August 1, 2016
Phase 5 (field interviewing)[1]	September 12, 2016	September 12, 2016
Phase 6 (prioritized data collection effort)[1]	November 17, 2016	November 17, 2016
Phase 7 (abbreviated interview)[1]	December 12, 2016	December 12, 2016
End of data collection[1]	January 31, 2017	January 31, 2017

[1] Beginning with phase 5, calibration sample and main sample cases were combined for data collection treatments.
NOTE: Subgroup A = high-school noncompleters; subgroup B = ultra-cooperative respondents; subgroup C = high school completers and sample members with unknown high school completion status.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

4.2.1 Data Collection Design Details

The HSLS:09 second follow-up data collection design incorporated the following features:

- three sample subgroups of interest for which interventions could be customized;

- seven distinct phases of data collection with corresponding interventions;
- a model to predict likelihood of response to maximize efficient allocation of project resources;
- a model to predict likelihood of contributing to nonresponse bias, used to identify sample cases for targeted interventions; and
- a calibration sample, fielded approximately 8 weeks in advance of the main sample to test the effectiveness of planned interventions experimentally.

This section describes the subgroups of interest in the HSLS:09 cohort, the response likelihood and bias likelihood models used to identify cases for targeted treatments, the phases of data collection, and the experimental calibration sample.

4.2.1.1 Subgroups of interest

The second follow-up sample was divided into three subgroups of interest, based on prior experience with the cohort, so that customized interventions could be developed based on patterns of response behavior from prior data collection rounds, and applied to each group independently. The subgroups consisted of the following:

1. **Subgroup A**—high school late/alternative/noncompleters (HSNC)—contained the subset of sample members who, as of the 2013 Update, had not completed high school, were still enrolled in high school, received an alternative credential, completed high school late, or experienced a dropout episode with unknown completion status.

2. **Subgroup B**—ultra-cooperative respondents (UC)—consisted of sample members who participated in the base year, first follow-up, and 2013 Update without an incentive offer. These cases were also early web respondents to the 2013 Update and are, by definition, high school completers.[16]

3. **Subgroup C**—high school completers and unknown high school completion status (HS other)—included cases that, as of the 2013 Update, were known to be on-time or early regular diploma completers not identified as ultra-cooperative, and cases with unknown high school completion status who were not previously identified as ever having had a dropout episode.

[16] In the spirit of a responsive design, the set of cases to be treated as "ultra-cooperative" was expanded for the main sample (i.e., cases not in the calibration sample) with the goal of maximizing the efficient use of project resources because response rates were relatively high. See section 4.2.1.6 for further details and for the expanded definition. The definition provided above corresponds to that used for sample members in the calibration sample.

4.2.1.2 Response likelihood model

Prior to data collection, a model was developed to predict the likelihood of a case responding to the survey: the response likelihood model. The model incorporated primarily paradata covariates from prior rounds—sampling frame information; base-year, first follow-up, and panel maintenance response outcomes; and 2013 Update survey data and paradata—that were found to be correlated with response outcomes. Table F-1 in appendix F provides a complete listing of response likelihood model covariates considered and retained in the final model. Using these prior data, a logistic regression model was fit predicting the response outcome in the 2013 Update as the dependent variable. Predicted probabilities were then used as a priori estimates of the probability that a sample member would respond in the second follow-up. The response likelihood model was run only once, before the start of data collection, and a response likelihood score was calculated for each sample member. During data collection, the response likelihood scores were used to assist in determining intervention resource allocation in phases 5 and 6 to avoid pursuing cases in field interviewing that were unlikely to respond; see section 4.2.1.4, below. See appendix F for details on the development of the response likelihood model.

4.2.1.3 Bias likelihood model

To estimate representativeness at various points during data collection, project staff estimated the variance between the weighted responding sample and baseline estimates of the overall population using a bias likelihood model run prior to intervention phases. The bias likelihood model was run at the beginning of phases 3 and 4 for the calibration samples (i.e., prior to each intervention) and at the beginning of phases 3, 4, 5, and 6 for the main sample cases. Note that modeling was done on the full sample (i.e., the combined calibration cases and main sample cases) after phase 4.

As was done for the responsive design approach during the 2013 Update data collection, the bias likelihood model used key survey and frame variables drawn from base-year, first follow-up, and 2013 Update survey data; High School Transcript data; school characteristics; and sampling frame information as predictors to identify nonrespondents who, unless converted to respondents, were most likely to contribute to bias in key survey variables. To calculate bias likelihood, a logistic regression model estimated the second follow-up response outcome and the predicted nonresponse probabilities derived from the model were then used to assign a bias likelihood score to each sample member. The bias likelihood score was defined as the predicted nonresponse probability output by the bias likelihood model. Details on the development and specifications of the model are provided in section F.2.2 in

appendix F. Nonresponding cases were then ordered according to the magnitude of their difference from the responding sample using the bias likelihood score, and cases most different (i.e., those most likely to contribute to bias if they remained nonrespondents) were targeted with interventions. Within phases 3 and 4, cases with bias likelihood scores above the median overall score were selected for targeting. Within phases 5 and 6, a cost-sensitive targeting approach was used in which the bias likelihood cutoffs were varied to meet in-phase targeting goals; see section 4.2.1.4 for further details. For details on the development of the bias likelihood model, see appendix F.

4.2.1.4 Data collection phases

As previously mentioned, the data collection design for the second follow-up included a responsive design with seven intervention phases. These phases included specific protocols for handling each of the three subgroups of sample members to reduce the potential for biased survey estimates or reduce data collection costs (Peytchev 2013). Results from the calibration sample experiments were used to determine the incentive levels used for the three subgroups' baseline incentive and each of two subsequent incentive boosts.

Baseline incentive (phase 1). During this initial phase, sample members were invited to complete the online or telephone interview through a data collection announcement letter and e-mail; no outbound telephone prompting occurred during this phase. Baseline incentive amounts were tested experimentally in the calibration sample (see complete results in section F.3 of appendix F), and then an optimal amount was applied to the main sample cases (i.e., those not selected for the calibration sample). Dependent upon their respective subgroup assignments, most sample members were offered a response-contingent baseline incentive which was announced in the data collection communication materials. Baseline incentive amounts ranged from $0 to $50 for the calibration sample, and from $0 to $40 for the main sample. For the calibration sample, incentive offers amounts were randomly assigned within subgroups; for the main sample, incentives were assigned within subgroup based on the experiment results from the calibration sample. This method of incentive assignment was used for the baseline incentive and the subsequent incentive boost offers. Phase 1 began in March 2016 for the calibration sample and in May 2016 for the main sample.

Outbound CATI prompting (phase 2). During phase 2, data collection interviewers (DCIs) prompted sample members by phone to complete the interview. Sample members were still eligible to receive the same baseline incentive that was offered to them in phase 1, if applicable. For most sample members, phase 2 began

3 weeks after the start of data collection; however, some sample members began phase 2 earlier.[17] Phase 2 began in early April 2016 for most of the calibration sample and in late May 2016 for the main sample.

Incentive boosts (phases 3 and 4). Phases 3 and 4 introduced the use of the bias likelihood model discussed in section 4.2.1.3 to target cases for two separate incentive boosts (boost 1 and boost 2). At the start of each phase, a subset of nonresponding cases was targeted to receive an increased incentive offer, in addition to the incentive(s) offered previously. Cases were selected for each boost independently of one another, that is, a case targeted in phase 3 might or might not be selected for targeting in phase 4, depending on how its bias likelihood score shifted between the phases. Because of their high analytic importance, all subgroup A (HSNC) cases were targeted for both boost 1 and boost 2 incentives. The calibration sample used random assignment to test the two different incentive boosts (see complete results in appendix F); like the baseline incentive, the optimal incentive amount was applied to the remaining main sample cases. Boost 1 ranged from $10 to $25 for the calibration sample and from $10 to $15 for the main sample (differing by subgroup). Boost 2 ranged from $10 to $20 for the calibration sample and was $10 for the main sample (varying by subgroup).[18] Phase 3 began in early May 2016 for the calibration sample and in late June 2016 for the main sample. Phase 4 began in June 2016 for the calibration sample and August 2016 for the main sample.

Field interviewing (phase 5). To conduct field interviewing as efficiently as possible, the start of phase 5 marked the end of the 8-week lag between the calibration sample and main sample; for phases 5 and 6, the calibration sample and main sample were combined, and all nonrespondents were eligible to be selected for field interviewing at the same time. Cases that were *not* selected for phase 5 targeting continued in data collection using the same contacting procedures employed in phase 4: mail and e-mail reminders and prompting calls from telephone interviewers. Cases that were selected for phase 5 targeting continued to receive the same mail and e-mail reminders as nontargeted cases; however, they no longer received prompting calls from telephone interviewers. Instead, the case was assigned to a FI for field telephone and in-person prompting. Several factors were used to select cases for field interviewing:

---

[17] High school noncompleters (i.e., cases within subgroup A) are often difficult to reach, but are of high analytic importance. Sample members who were identified as high school noncompleters began phase 2 after 1 week of data collection, to allow additional time for DCIs to contact these cases.
[18] An additional adaptive incentive boost component (boost 2b) of $10 was offered to main sample subgroup A (HSNC) nonrespondents and main sample subgroup C (HS other) cases targeted for boost 2. Note that this was offered to remaining nonrespondents, in addition to the $10 boost 2 incentive offer. See section 4.2.1.5 for further details.

- the response likelihood and bias likelihood scores calculated for the case as part of the responsive design modeling approach;
- whether the case was part of a subgroup of interest;
- the availability of contact leads;
- the status of the case;
- proximity to a geographic cluster of other selected field cases; and
- the caseload of the FIs working in that geographic area.

Field cases were selected in multiple waves. The initial wave of cases was selected in early September 2016, with field data collection beginning in mid-September.

Due to the high relative cost of field interviewing compared to other data collection modes, the maximum size for the initial wave was set at approximately 2,000 cases. To exclude cases that had the lowest probability of responding, the minimum response likelihood score[19] threshold for field cases was set at .50; cases below .50 were not eligible to be selected for field data collection; note that overall response likelihood scores had a mean of .80 and standard deviation of .28. Cases were then prioritized for field interviewing based on their bias likelihood scores, with those cases having the highest likelihood of contributing to bias if they remained nonrespondents given higher priority. Priority was also given to subgroup A (HSNC) cases. Among subgroup A cases, the minimum bias likelihood score[20] threshold was set at .54; for all other cases, the bias likelihood threshold was set at .58; note that overall bias likelihood scores at phase 5 had a mean of .47 and standard deviation of .15. Cases that had been previously contacted during second follow-up data collection and refused to participate, and those with neither an address nor phone number on record, were excluded from the field selection. The selected cases were then assigned to the closest FI geographically.

As FIs' caseloads decreased, either because sample members completed the survey or all leads were exhausted, additional field cases were selected. Prior to selecting new field cases, the data collection team determined the number of additional cases that each interviewer should receive to maintain an optimal caseload. Available field cases were then assigned to geographic clusters based on the closest FI. Cases with no known address were grouped based on their current or most recently available

[19] The response likelihood score is a continuous probability of response, bounded by 0 and 1, with a value of 1 indicating a case is predicted to respond and 0 indicating a case is predicted not to respond.
[20] The bias likelihood score is bounded by 0 and 1, with a value of 1 indicating a nonresponding case is most likely to contribute to bias if remaining a nonrespondent and 0 indicating a case is least likely to contribute to bias if remaining a nonrespondent.

telephone area code.[21] For these later selection waves, the response likelihood threshold of .50 was retained, and the bias likelihood threshold was adjusted within each cluster until enough cases had been selected to meet the caseload goals. By the end of data collection, approximately 3,600 cases had been selected for field interviewing. Table 9 in section 4.3.2 provides the participation status of field cases, by interview mode.

Prioritized data collection effort (phase 6). Phase 6 involved allocating data collection effort to cases based on their importance, as determined using the two responsive design models. The purpose of phase 6 was to concentrate data collection effort on the most important cases among all pending nonrespondents.

Cases most likely to contribute to bias if they remained nonrespondents (i.e., those with the highest bias likelihood values) and most likely to respond based on a priori response likelihood score (i.e., those with the highest response likelihood values) were targeted, and cases that were least likely to contribute to bias if they remained nonrespondents (i.e., lowest bias likelihood values) and the least likely to respond based on response likelihood score (i.e., lowest response likelihood values) were untargeted. Targeted cases were those with a response likelihood score above .50 and a bias likelihood score above .40. Due to their analytical importance, all subgroup A (HSNC) cases were selected for targeting. All pending nonrespondent cases that had been targeted for field interviewing in phase 5 were also targeted for phase 6 prioritization.

Data collection protocols in phase 6 were the same as those used in phase 5, but phase 6 targeted cases were prioritized first for tracing and telephone prompting and were eligible for field interviewing. Untargeted cases received reduced data collection effort by being placed at the bottom of the tracing and calling queues. As FIs' caseloads decreased, additional cases were selected for field interviewing, using the procedures also used in phase 5. However, cases not targeted for phase 6 were not eligible to be selected for field interviewing. Data collection remained open to allow untargeted cases to participate, but efforts to pursue those cases were reduced.

A total of 6,303 cases, or 91 percent of pending nonrespondent cases, were targeted for phase 6. The remaining 9 percent of cases were untargeted for phase 6. Phase 6 began in November 2016.

[21] These cases without an address were worked by the field interviewer over the phone. If an address was later obtained for the sample member, either through field locating or through centralized intensive tracing, the case was transferred to the nearest field interviewer, as appropriate.

Abbreviated interview (phase 7). As discussed in section 3.3, approximately 7 weeks before the end of data collection, the full-length survey instrument was replaced with an abbreviated interview which was subsequently offered to all pending nonrespondents to increase the overall response rate and to encourage participation across the entire pending sample. The abbreviated interview could be completed in approximately half the time of the full-length interview (17 minutes versus 32 minutes; see section 4.3.3 for details on survey timing). The abbreviated interview was implemented in December 2016 and continued until the January 2017 end of data collection. Note that the abbreviated interview was not targeted to a subset of respondents; instead, all eligible pending nonrespondents were offered the option to complete the abbreviated interview. Sample members were notified of the abbreviated interview offer through every contact method available: a reminder postcard mailing for those sample members with a mailing address, e-mail for those sample members with an e-mail address, over the telephone by a DCI for those sample members with a phone number, and by an FI for those sample members selected for field interviewing. Protocols which began in phase 6, in which targeted cases received prioritized data collection effort, continued unchanged after the abbreviated interview was offered.

4.2.1.5 Calibration sample

A randomly selected calibration sample of 3,300—approximately 14 percent of the second follow-up fielded sample—was fielded in advance of the main sample and was used to evaluate optimal incentive interventions within each of the three subgroups. Table 4 shows the sample size of each subgroup and the number of cases selected for the calibration sample as well as the main sample.

Table 4. Calibration sample and main sample sizes, by subgroup: 2016

Subgroup	Second follow-up fielded sample	Calibration sample	Main sample
Total	23,316	3,300	20,016
Subgroup A (high school late/alternative/noncompleters)	2,545	663	1,882
Subgroup B (ultra-cooperative respondents)	4,144	663	3,481
Subgroup C (all other high school completers and unknown cases)	16,627	1,974	14,653

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

Both the calibration sample and the main sample were subject to the same data collection activities, although the calibration sample was fielded approximately

8 weeks before the main sample. After the completion of phase 4 (described above in section 4.2.1.4), pending nonrespondents from the calibration sample and the main sample were combined for the remainder of data collection.

Three incentive experiments were embedded in the data collection for the calibration sample to inform the incentive amounts in the subsequent data collection of the main sample. The first experiment identified the baseline incentive, and the second and third determined two subsequent incentive increases, or boosts. These activities, which occurred at different points during data collection, are summarized below. For further details on the calibration sample experiments and their results, see appendix F.

Baseline incentive (phases 1 and 2). To assess how baseline incentive offers should vary by subgroup, calibration sample members were randomly assigned to baseline incentive amounts of $0, $30, $40, and $50 for subgroup A (HSNC) and subgroup B (UC), and amounts ranging from $15 to $40, in $5 increments, for subgroup C (HS other).

At the end of phase 2, results were analyzed to inform the main sample data collection. An incentive of $40 was determined to be optimal for subgroup A (HSNC). For subgroup B (UC), the lack of an incentive offer (i.e., $0) was found to elicit a relatively high response rate (64 percent) when measured against comparable subgroups in BPS:12/14 (those with high response likelihood scores) and B&B:08/12 (double respondents and early respondents). Finally, an incentive of $30 was identified as most appropriate for subgroup C (HS other), based on the experiment results and prior experiences with similar populations.

Incentive boost 1 (phase 3). The second experiment was implemented in phase 3 and was designed to determine the additional incentive amount to offer to select main sample nonrespondents. This experiment introduced the use of the bias likelihood model, described in section 4.2.1.3, to target cases for intervention. Cases identified for targeting were randomized to an incentive boost in addition to the baseline incentive offer. All subgroup A (HSNC) cases were targeted, due to this group's importance, and were subsequently randomized to boosts of $15 and $25. Targeted cases from both subgroup B (UC) and subgroup C (HS other) were randomized to the $10 and $20 boost categories.

At the end of phase 3, HSLS:09 staff evaluated the experiment results to inform the main sample collection. Among subgroup A (HSNC) cases, an incentive of $15 was found to be optimal. For subgroup B (UC) cases, a boost incentive of $10 was

identified as most appropriate for targeted cases. A boost incentive of $10 was also chosen for subgroup C (HS other) cases.

Incentive boost 2 (phase 4). In this third experiment, all remaining nonrespondent cases that were targeted based on the results of the bias likelihood model were randomized to incentive boosts of either $10 or $20. Four weeks after the initiation of the boost 2 incentive, response rates were compared to inform main sample data collection. No statistical difference in response rates between the $10 and $20 groups was observed at that time. As a result, $10 was identified as the optimal amount for all subgroups for the main sample boost 2 amount.

Nonresponse follow-up incentive. Approximately 11 weeks after the start of phase 4 (implementation of incentive boost 2), the response rate for calibration subgroup A (HSNC) cases that did not receive a baseline incentive offer (28 percent) lagged far behind the response rate for calibration subgroup A cases that did receive an offer (47 percent). To encourage participation and to reduce the significant response rate gap, an additional incentive of $40 was offered approximately 14 weeks after the implementation of incentive boost 2 to pending calibration HSNC cases that did not receive a baseline incentive offer.

4.2.1.6 Calibration-informed adaptive components

Two other adaptive components, informed by findings in the calibration sample, were implemented for the main sample: an expanded definition for subgroup B (ultra-cooperative cases) and an additional incentive boost 2 (2b).

Expanded definition of ultra-cooperative cases. As noted above in section 4.2.1.5, the response rate for the ultra-cooperative calibration sample at the end of phase 2 was relatively high, at 64 percent, for the sample members who were not offered a monetary incentive. To broaden this set to include more cases as a cost-sensitive and cost-containment strategy, thereby increasing the number of cases not offered a baseline incentive, the definition for the main sample ultra-cooperative cases was expanded to also include sample members satisfying all the following criteria:

- must have been in subgroup C (HS other);
- must have predicted response likelihood > .90;
- must have been an early/on-time high school completer;
- must not have ever dropped out of high school;
- must have been a 2013 Update respondent;
- in 2013 Update, the respondent must have been the student not the parent;
- in 2013 Update, must have responded in the first 12 weeks; and
- in 2013 Update, must not have received a $5 prepaid incentive.

The original set of main sample subgroup B (UC) cases included 364 high school completers who participated in the base year and first follow-up and completed the 2013 Update in the early web period, with no incentive. According to the expanded definition and revised criteria, an additional 3,117 cases were reclassified from subgroup C (HS other) to subgroup B (UC) for the main sample. Note that table 4, above, reflects the number of cases in each subgroup following the reclassification.

Incentive boost 2b. When boost 2 response rates were reassessed 11 weeks after the start of phase 4, the differences among calibration sample cases had become large and statistically significant for subgroups A (HSNC) and C (HS other) cases. Subgroup B (UC) had very small numbers and no observed difference. Based on these findings, an additional adaptive boost (incentive boost 2b) of $10 was offered to main sample subgroup A (HSNC) nonrespondents and main sample subgroup C (HS other) cases targeted for boost 2. Sample members were notified of the increased incentive offer by e-mail, by a telephone interviewer (for cases in outbound CATI prompting), or by an FI (for cases selected for field interviewing). The increased incentive offer amount was also communicated in the next reminder mailing.

4.3 Data Collection Results

The HSLS:09 project team assessed data collection outcomes by reviewing the number of sample members located and interviewed, the interview mode selected by respondents, the number of sample members that required intensive tracing, the conversion of interview refusals, and survey response timing.

For purposes of reporting second follow-up data collection results, the HSLS:09 sample can be separated into four categories:

1. *Ineligible* sample members include sample members who are deceased or found to be study ineligible prior to the start of second follow-up data collection.[22]
2. A first set of *out-of-scope* cases consists of sample members determined to be out of scope prior to the start of second follow-up data collection. This set includes sample members who had withdrawn from the study and requested that they not be re-contacted, and sample members who were nonrespondents both in the base year and first follow-up.

[22] Sample members were classified as study ineligible if they were not in 9th grade during the base-year data collection, if they were not enrolled at the sampled high school during the base year, or if they were foreign exchange students.

3. A second set of *out-of-scope* cases includes fielded sample members determined during the second follow-up data collection to be out of the country, incarcerated, institutionalized, deceased, or otherwise excluded from the second follow-up interview because it was not offered in a format that allowed their meaningful participation.
4. Sample members who are not in one of the three previous sets are defined to be *in-scope* for data collection.

Fielded sample members correspond to sets 3 and 4, above. Cases determined to be out of scope during the second follow-up data collection—corresponding to set 3—were excluded from further data collection effort after determining their disposition. This section reports on the degree to which data collection efforts yielded participants among set 4, the in-scope set of fielded sample members.

In calculating response rates, there are two types to consider: *participation rate* and *response rate*. The participation rate removes from the denominator all categories of ineligible and out-of-scope sample members, corresponding to sets 1 through 3, above. The participation rate is a measure of the methodological success of the data collection round, conveying what the data collection effort accomplished among the set of cases pursued. Participation rates are calculated using unweighted data. The weighted response rate is the rate of response calculated with exclusions made only for previously identified deceased and study ineligible sample members—corresponding to set 1, above—and cases found to be deceased as of the second follow-up, which are removed from the study denominator. Thus, the response rate reflects the proportion of the eligible target population[23] represented by sample respondents, and therefore serves as an indicator of data quality. Higher response rates can be an indicator of more accurate survey results, where the responding sample better represents the target population of interest. However, it is also important to examine the potential for nonresponse bias. Chapter 6 includes a discussion of study response rates and presents information on potential nonresponse bias for unit nonresponse. Chapter 6 and appendix G also provide item nonresponse bias information for weighted item response rates that fall below 85 percent.

Table 5 shows participation rates, by prior response status and student type, for the second follow-up. Among the fielded sample that were in scope for data collection, 17,335 sample members participated in the second follow-up, resulting in a

[23] The HSLS:09 target population is all students in the 9th grade during the fall 2009 term who attended either regular public or private schools in the 50 states and the District of Columbia that provided instruction in both 9th and 11th grades. See chapter 2 for additional information on the study samples.

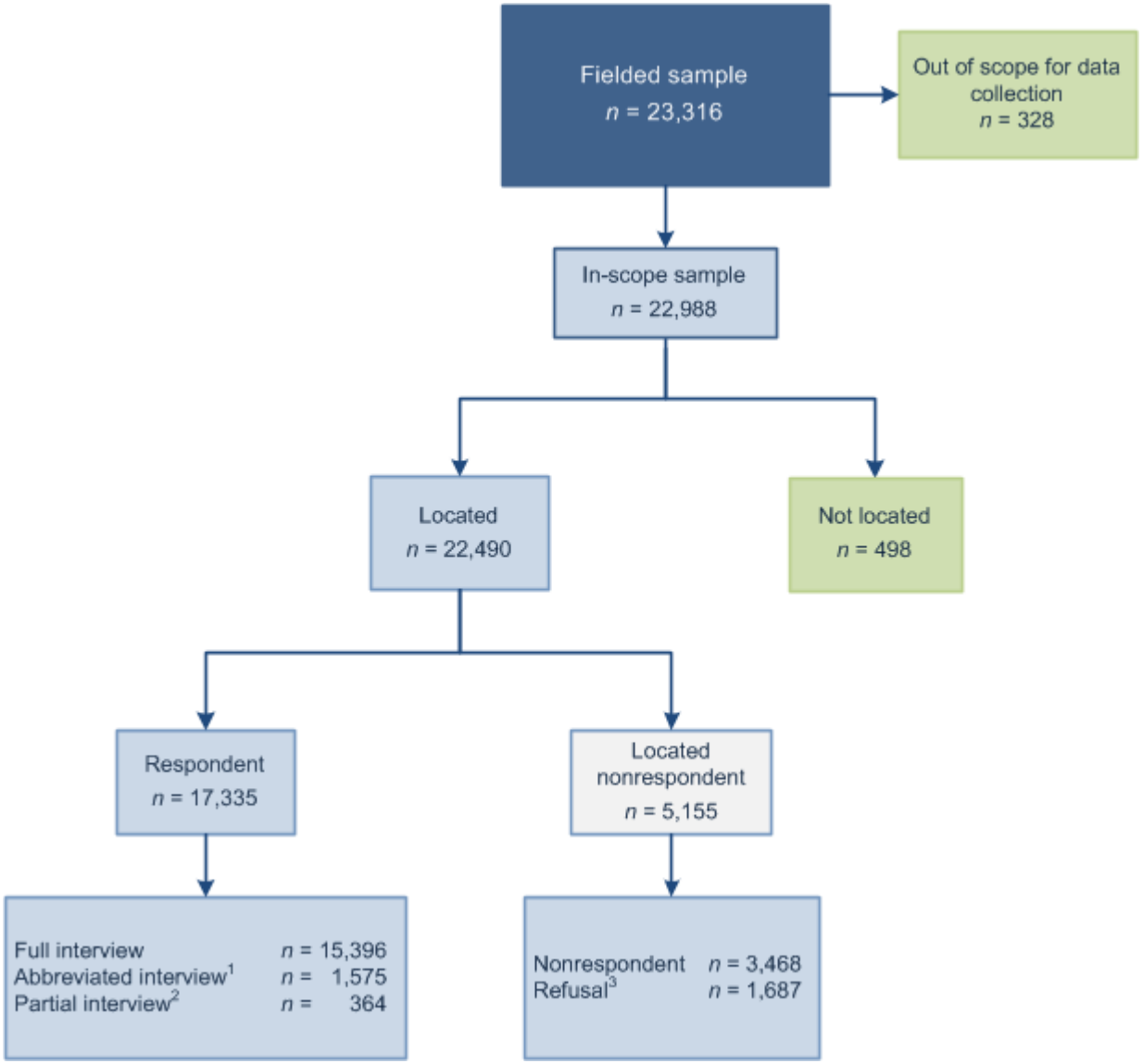
participation rate of approximately 75 percent. Figure 8 shows the overall data collection outcomes for the fielded second follow-up sample. Of the fielded sample, 328 cases were determined to be out of scope for data collection because they were deceased, incarcerated, or otherwise excluded from the second follow-up interview because it was not offered in a format that allowed their meaningful participation; these cases were excluded from further data collection effort. Of the 22,988 in-scope cases, 498 were not located, 22,490 were located, and 17,335 responded.

Table 5. Participation rates, by prior response status and student type: 2016

Prior response status and student type	Total fielded sample	In-scope for data collection[1]	Participated	
			Number	Percent of eligible
Total	**23,316**	**22,988**	**17,335**	**75.4**
Panel maintenance response status (2015)				
Panel maintenance respondent	7,956	7,918	7,418	93.7
Panel maintenance nonrespondent	15,360	15,070	9,917	65.8
2013 Update response status				
Respondent	18,556	18,349	15,242	83.1
Nonrespondent	4,760	4,639	2,093	45.1
First follow-up response status				
Respondent	20,520	20,276	16,106	79.4
Nonrespondent	2,796	2,712	1,229	45.3
High school completion status[2]				
High school diploma	19,295	19,065	15,127	79.3
Alternative credential[3]	629	620	455	73.4
High school noncompleter[4]	1,481	1,459	1,097	75.2
Unknown status	1,911	1,844	656	35.6

[1] Excludes 328 cases that were part of the initial fielded sample but were excluded from further data collection effort because they were found to be deceased, incarcerated, or otherwise excluded from the second follow-up interview because it was not offered in a format that allowed their meaningful participation.
[2] High school completion status from any source, prior to the start of second follow-up data collection.
[3] Includes cases that received a General Educational Development (GED) or other alternative high school credential.
[4] Includes all cases that did not have a high school credential, prior to the start of second follow-up data collection (including those that were still enrolled in a high school completion program).
NOTE: Detail may not sum to totals because of rounding. The participation count includes sample members who met the criteria for qualification as an interview respondent, which required completing at least a partial interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Figure 8. Overall locating and interviewing results: 2016



[1] Includes 148 cases that began the full interview and completed the abbreviated interview.
[2] Cases completed enough of the full or abbreviated interview to be considered a respondent but did not complete the entire interview. See section 3.5 for a description of the rules used to define interview respondents.
[3] Includes all cases that ended data collection at a refusal status, regardless of the source of the refusal; the sample member may have refused, or another contact may have refused on the sample member's behalf.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

4.3.1 Tracing and Locating Outcomes

Tracing and locating rates varied by high school completion status, prior interview response status, and panel maintenance response status. As a result of all tracing activities and contact attempts throughout the data collection period, approximately

96 percent of cases were located. Table 6 presents located rates for the fielded second follow-up sample.

Table 6. Located status, by prior response status and student type: 2016

Prior response status and student type	Total fielded sample	In-scope for data collection[1]	Located	
			Number	Percent of in-scope
Total	**23,316**	**22,988**	**22,490**	**97.8**
Panel maintenance response status (2015)				
Panel maintenance respondent	7,956	7,918	7,913	99.9
Panel maintenance nonrespondent	15,360	15,070	14,577	96.7
2013 Update response status				
Respondent	18,556	18,349	18,349	100.0
Nonrespondent	4,760	4,639	4,141	89.3
First follow-up response status				
Respondent	20,520	20,276	20,053	98.9
Nonrespondent	2,796	2,712	2,437	89.9
High school completion status[2]				
High school diploma	19,295	19,065	18,906	99.2
Alternative credential[3]	629	620	609	98.2
High school noncompleter[4]	1,481	1,459	1,449	99.3
Unknown status	1,911	1,844	1,526	82.8

[1] Excludes 328 cases that were part of the initial fielded sample but were excluded from further data collection effort because they were found to be deceased, incarcerated, or otherwise excluded from the second follow-up interview because it was not offered in a format that allowed their meaningful participation.
[2] High school completion status from any source, prior to the start of second follow-up data collection.
[3] Includes cases that received a General Educational Development (GED) or other alternative high school credential.
[4] Includes all cases that did not have a high school credential, prior to the start of second follow-up data collection (including those that were still enrolled in a high school completion program).
NOTE: Detail may not sum to totals because of rounding. Sample members are counted as located if they were ever located during data collection.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Batch tracing. HSLS:09 staff submitted all existing contact information to the NCOA database. Of the 23,253 cases with addresses sent to NCOA, new or confirmed contact information was returned for 17,138 (74 percent). Phone Append returned new or confirmed telephone numbers for 1,944 (57 percent) of the 3,387 cases sent. Before intensive tracing, a small group of cases was sent to Premium Phone because all telephone leads had been exhausted during outbound dialing. Of the 143 cases sent to Premium Phone, 16 (11 percent) were matched with a new telephone number. Table 7 lists batch tracing results for NCOA, Phone Append, Premium Phone, and other batch tracing sources used for HSLS:09.

Table 7. Batch processing case match rates, by tracing source: 2016

Method of tracing	Number of cases sent	Number of cases matched	Percent matched
National Change of Address	23,253	17,138	73.7
LexisNexis Phone Append	3,387	1,944	57.4
First Data Premium Phone	143	16	11.2
LexisNexis SSN Search	4,520	261	5.8
Experian TrueTrace	14,611	10,091	69.1
LexisNexis Single Best Address	3,907	3,564	91.2
LexisNexis Single Best Phone	2,949	1,052	35.7
LexisNexis E-mail Search	2,696	443	16.4

NOTE: Match rate includes instances when sample member contact information was confirmed and when new information was provided. Multiple records per case may have been sent to each source, so the actual number of records matched is higher than the number of unique cases matched.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Intensive tracing. The intensive tracing process used known identifying information (e.g., date of birth, SSN, and previous address information) to search for a sample member through credit reports and other private-use databases. Table 8 shows results for cases requiring intensive tracing and results for those for whom new locating information was obtained through intensive tracing, by panel maintenance response status, prior interview response status, and high school completion status. Cases with unknown high school completion status required intensive tracing at a higher rate than any other subgroup, with 29 percent of those cases undergoing tracing efforts. Panel maintenance respondents required intensive tracing at the lowest rate, with only 1 percent of those cases pursued through intensive tracing. Among cases that required intensive tracing, new locating information was obtained through intensive tracing at rates that ranged from 25 percent for cases with unknown high school completion status to 31 percent for cases with a high school diploma.

Table 8. Intensive tracing rates and rates located though intensive tracing, by prior response status and student type: 2016

Prior response status and student type	In-scope sample[2]	Required intensive tracing		Located through intensive tracing[1]	
		Number	Percent of in-scope	Number	Percent of traced
Total	**22,988**	**1,829**	**8.0**	**532**	**29.1**
Panel maintenance response status (2015)					
Panel maintenance respondent	7,918	107	1.4	31	29.0
Panel maintenance nonrespondent	15,070	1,722	11.4	501	29.1
2013 Update response status					
Respondent	18,349	901	4.9	260	28.9
Nonrespondent	4,639	928	20.0	272	29.3
First follow-up response status					
Respondent	20,276	1,217	6.0	369	30.3
Nonrespondent	2,712	612	22.6	163	26.6
High school completion status[3]					
High school diploma	19,065	920	4.8	283	30.8
Alternative credential[4]	620	116	18.7	38	32.8
High school noncompleter[5]	1,459	260	17.8	76	29.2
Unknown status	1,844	533	28.9	135	25.3

[1] Indicates new and confirmed locating information was obtained through the intensive tracing process.
[2] Excludes 328 cases that were part of the initial fielded sample but were excluded from further data collection effort because they were found to be deceased, incarcerated, or otherwise excluded from the second follow-up interview because it was not offered in a format that allowed their meaningful participation.
[3] High school completion status from any source, prior to the start of second follow-up data collection.
[4] Includes cases that received a General Educational Development (GED) or other alternative high school credential.
[5] Includes all cases that did not have a high school credential, prior to the start of second follow-up data collection (including those that were still enrolled in a high school completion program).
NOTE: Detail may not sum to totals because of rounding. Total excludes cases initiated to intensive tracing that were not traced.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

4.3.2 Interview Participation Rates

Participation by mode. As described in section 3.4, second follow-up interviews were administered in several modes, including web (mobile and nonmobile devices), telephone, and in person. Table 9 shows the distribution of second follow-up respondents by interview mode. Sample members most often chose to complete the self-administered web interview, with 80 percent of respondents choosing this option; 29 percent of all respondents completed on a mobile device; and 50 percent of respondents completed on a nonmobile device. Fifteen percent of respondents completed the interview with a centralized telephone interviewer (i.e., DCI). Approximately 5 percent of all interviews were completed with an FI, whether in

person (1 percent), over the telephone (3 percent), or through self-administration in the field (less than 1 percent).

Table 9. Distribution of interview participation, by interview mode: 2016

Interview mode	Total	
	Number	Percent of all respondents
Total	**17,335**	**100.0**
All web interviews	13,836	79.8
Web (mobile)	5,092	29.4
Web (nonmobile)	8,744	50.4
Centralized telephone[1]	2,613	15.1
Field in person	256	1.5
Field telephone[2]	576	3.3
Field self-administered	54	0.3

[1] Interview was conducted with a data collection interviewer (either from inbound or outbound call).
[2] Interview was conducted over the telephone by a field interviewer.
NOTE: Detail may not sum to totals because of rounding. Interviewed count includes eligible sample members who met the criteria for qualification as an interview respondent, which required completing at least a partial interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Field data collection. The first wave of field cases was selected in September 2016, and two additional waves were selected in October and November 2016. By the end of data collection, 3,592 cases had been selected for field interviewing, and approximately 50 percent of those cases completed the interview. When cases were selected for field interviewing, they were transferred from centralized outbound calling to FIs to work over the telephone or in person. Although no further outbound CATI calls were made to field cases, the cases selected for field interviewing could still call the help desk and complete the interview over the telephone with a DCI. Approximately 5 percent of all field respondents completed a telephone interview with a DCI. Field cases could also complete the self-administered web interview; of field respondents, 46 percent completed the self-administered web interview. Approximately 14 percent of field respondents completed the interview in person with an FI, 32 percent over telephone with an FI, and 3 percent self-administered the interview in the field.

Table 10 shows the response status for all second follow-up field cases.

Table 10. Participation status of field cases, by interview mode: 2016

Interview mode	Number of field cases	Percent of all field cases	Percent of all field respondents
Total	**3,592**	**100.0**	**†**
Respondents	1,787	49.7	100.0
All web interviews	817	22.7	45.7
Web (mobile)	478	13.3	26.7
Web (nonmobile)	339	9.4	19.0
Centralized telephone[1]	84	2.3	4.7
Field in person	256	7.1	14.3
Field telephone[2]	576	16.0	32.2
Field self-administered	54	1.5	3.0
Nonrespondents	1,805	50.3	†

† Not applicable.
[1] Interview was conducted with a data collection interviewer.
[2] Interview was conducted over the telephone by a field interviewer.
NOTE: Detail may not sum to totals because of rounding. Respondent count includes eligible sample members who met the criteria for qualification as an interview respondent, which required completing at least a partial interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Completion by phase of data collection. As described in section 4.2.1.4, the second follow-up data collection was conducted in seven phases, corresponding to responsive design interventions intended to reduce nonresponse bias among underrepresented groups. Table 11 shows the distribution of interview respondents by data collection phase. The greatest number of respondents completed during phase 1 (early web-only data collection), with 31 percent of respondents completing during this phase; the fewest responded during phase 6 (prioritized data collection effort), with 5 percent of respondents completing during this phase. See appendix F for an evaluation of the responsive design interventions.

Table 11. Distribution of interview respondents, by data collection phase: 2016

Data collection phase	Total	
	Number	Percent of all respondents
Total	**16,971**	**100.0**
Phase 1: Early web	5,237	30.9
Phase 2: Outbound CATI	2,750	16.2
Phase 3: Incentive boost 1	2,786	16.4
Phase 4: Incentive boost 2	2,074	12.2
Phase 5: Field interviewing	1,759	10.4
Phase 6: Prioritized data collection effort	761	4.5
Phase 7: Abbreviated interview	1,604	9.5

NOTE: Detail may not sum to totals because of rounding. Includes respondents who completed a full or abbreviated interview; excludes 364 partial interview respondents. By definition, partial interview respondents never fully completed the interview in a single phase and therefore stayed active for subsequent phases. As such, these cases are excluded from this table because their partial response cannot be attributed to a single phase.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Interview completeness. A total of 15,396 sample members completed the full second follow-up interview. In December 2016, all pending nonrespondents were offered an abbreviated interview; 1,575 sample members, or 9 percent of all respondents, completed the abbreviated interview. An additional 364 cases completed enough of either the full or abbreviated interview to be categorized as a partial interview respondent. Section 3.5 provides a description of the rules used to define interview respondents. Table 12 shows interview completeness for second follow-up respondents.

Table 12. Interview completeness among respondents: 2016

Interview completeness	Number of respondents	Percent of all respondents
Total	**17,335**	**100.0**
Full interview	15,396	88.8
Abbreviated interview	1,575	9.1
Partial interview	364	2.1

NOTE: Numbers may not sum to total due to rounding. Respondent count includes sample members who met the criteria for qualification as an interview respondent, which required completing at least a partial interview. Sample members that began the full interview but later completed the abbreviated instrument are included as abbreviated interview cases. Excludes sample members who did not complete enough of the interview to qualify as an interview respondent.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Refusal conversion. As noted in section 4.1.3, second follow-up staff integrated refusal conversion techniques into DCI training and reinforced them throughout data collection in QC meetings. Interviewers were encouraged to share their

experiences in avoiding sample member refusals and to seek guidance from PTLs for particularly difficult cases. Sample member refusals were classified by strength of refusal—from "soft" refusals ("I'm just not interested") to hostile, firm refusals. Hostile refusals were immediately finalized, and soft refusals were called back for conversion efforts. Project staff placed sample members who refused to complete the interview in a separate calling queue that was staffed by a subset of interviewers who had received specialized refusal conversion training. Overall, around 4 percent of in-scope cases ever refused; of those, about 16 percent of cases subsequently completed the interview (table 13).

Table 13. Refusal and refusal conversion rates, by prior response status and student type: 2016

Prior response status and student type	In-scope sample[1]	Ever refused interview[2]		Interviewed, of ever refused	
		Number	Percent of eligible	Number	Percent of refusals
Total	**22,988**	**942**	**4.1**	**147**	**15.6**
Panel maintenance response status (2015)					
Panel maintenance respondent	7,918	157	2.0	44	28.0
Panel maintenance nonrespondent	15,070	785	5.2	103	13.1
2013 Update response status					
Respondent	18,349	565	3.1	110	19.5
Nonrespondent	4,639	377	8.1	37	9.8
First follow-up response status					
Respondent	20,276	730	3.6	124	17.0
Nonrespondent	2,712	212	7.8	23	10.8
High school completion status					
High school diploma[3]	19,065	777	4.1	124	16.0
Alternative credential[4]	620	25	4.0	6	24.0
High school noncompleter[5]	1,459	33	2.3	8	24.2
Unknown status	1,844	107	5.8	9	8.4

[1] Excludes 328 cases that were part of the initial fielded sample but were excluded from further data collection effort because they were found to be deceased, incarcerated, or otherwise excluded from the second follow-up interview because it was not offered in a format that allowed their meaningful participation.
[2] Includes sample members who refused to participate in the second follow-up interview; excludes refusals if another contact (such as a parent) refused on the sample member's behalf.
[3] High school completion status from any source, prior to the start of second follow-up data collection.
[4] Includes cases that received a General Educational Development (GED) or other alternative high school credential.
[5] Includes all cases that did not have a high school credential prior to the start of second follow-up data collection (including those that were still enrolled in a high school completion program).
NOTE: Interviewed count includes sample members who met the criteria for qualification as an interview respondent, which required completing at least a partial interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Incentives. Sample members who completed the telephone or self-administered web interview could opt to receive their incentive payment by paper check or through PayPal. Among sample members offered an incentive, approximately 68 percent of respondents chose a paper check, 31 percent chose PayPal, and 1 percent declined the incentive or did not make a selection. Table 14 shows the proportion of respondents who selected each incentive option. Respondents who completed the interview with an FI, whether in person or over the phone, only had one incentive option (cash for in-person field interviews and paper check for telephone field interviews); therefore, field respondents are not included in table 14.

Table 14. Incentive selection among web and centralized telephone respondents who were offered an incentive: 2016

	Total	
Incentive selection	Number	Percent of respondents
Total web and centralized telephone respondents	**14,514**	**100.0**
Check	9,922	68.4
PayPal	4,451	30.7
Declined/No selection[1]	141	1.0

[1] Includes sample members who declined the incentive and those who exited the survey instrument without making a selection.
NOTE: Detail may not sum to totals because of rounding. Includes only cases that completed enough of the self-administered web interview or a centralized telephone (CATI) interview to qualify as a respondent. Excludes cases that completed a field interview, either in person or over the telephone as these respondents had only a single incentive option (cash for in-person field interviews and paper check for telephone field interviews). Also excludes cases that completed an interview but were not offered an incentive.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

4.3.3 Survey Timing

To assess the burden associated with completing the second follow-up survey, the time required for each respondent to complete the interview was collected and analyzed in aggregate. Both the full-length interview and abbreviated interview were assessed separately by survey mode of completion, overall and by section.

To calculate the total instrument time, form-level times were summed across the survey. Likewise, section-level times were computed by summing across all forms within a given section. Individual form times were computed using a time-stamp embedded on each form in the interview. A timer recorded the clock time on a respondent's or interviewer's computer when a form was first loaded to obtain the start time on that form, and an end timer recorded the clock time when the "Next" button on the form was clicked to calculate the end time on the form. Form-level times were then calculated by subtracting the start time from the end time. Additionally, outliers were identified using the interquartile range (IQR) and defined

as $Q_3 + (5 \cdot \text{IQR})$ at the form level, where Q_3 is the third quartile. A lower bound was not defined for outliers. Outliers were recoded to the median value at the form level to allow for aggregation of form times to create section times and overall survey times for all respondents. Cases that began the full-length interview but later completed the abbreviated instrument are excluded from analyses.

4.3.3.1 Full instrument

Full instrument overall timing. Overall, the full-length survey took 31.6 minutes to complete on average. Compared to telephone interviews[24] (averaging 42.7 minutes), self-administered interviews completed via web mode (which includes those completed on both mobile and nonmobile devices) took significantly less time to complete ($t(3798.67) = 70.22$, $p < .05$), and averaged 29.1 minutes. No statistically significant timing difference was found between self-administered mobile interview (averaging 28.9 minutes) and nonmobile interview (averaging 29.2 minutes) modes, indicating that the survey was not overly burdensome to complete on devices with reduced screen-size, relative to nonmobile devices. The field interview[25] averaged 37.8 minutes to complete, which was significantly longer than self-administered interviews completed via the Web ($t(220.68) = 12.37$, $p < .05$). The significant differences in timing observed between self-administered web surveys and telephone as well as field interviews is expected, given the additional time required for FIs to read questions and other text aloud to respondents.

Full instrument section timing. Average section completion times were 1.2 minutes for the High School section, 10.2 minutes for Postsecondary Education, 8.8 minutes for Employment, 5.9 for Family and Community, and 5.1 for the Locating section.

Self-administered web surveys took significantly less time than telephone interviews for all sections. Among self-administered web surveys and field interviews, web surveys took less time to complete for all sections, except the Postsecondary Education section which was found to be significantly longer for web surveys ($t(219.14) = 2.68$, $p < .05$). Table 15 shows the average time in minutes to complete the full-length interview, by interview section and mode.

[24] For timing analyses, "telephone interview" includes cases completed through centralized telephone interviewing and field telephone interviewing.
[25] For timing analyses, "field interview" includes cases completed through field in-person (CAPI) and field self-administration.

Table 15. Average time in minutes to complete the full interview, by interview section and mode: 2016

			Mode of administration									
	All modes		Web (nonmobile)		Web (mobile)		Web[1]		Telephone[2]		Field[3]	
Interview section	Number	Time	Number	Time	Number	Time	Number	Time	Number	Time	Number	Time
All sections[4]	**15,396**	**31.6**	**8,153**	**29.2**	**4,389**	**28.9**	**12,542**	**29.1**	**2,638**	**42.7**	**216**	**37.8**
High School	15,392	1.2	8,152	1.0	4,389	1.2	12,541	1.1	2,635	1.7	216	2.5
Postsecondary Education	15,394	10.2	8,152	10.1	4,389	9.1	12,541	9.8	2,637	12.7	216	8.6
Employment	15,395	8.8	8,153	8.0	4,389	8.2	12,542	8.1	2,637	12.0	216	11.9
Family and Community	15,396	5.9	8,153	5.1	4,389	5.4	12,542	5.2	2,638	9.2	216	8.4
Locating	15,396	5.1	8,153	4.6	4,389	4.7	12,542	4.7	2,638	7.1	216	5.9

[1] Includes self-administered nonmobile and mobile interviews.
[2] Includes centralized telephone and field telephone interviews.
[3] Includes field in-person (CAPI) and field self-administered cases.
[4] Includes timing for the Consent section, which preceded the interview sections.
NOTE: Detail may not sum to totals because of rounding. Partial interviews and cases that began the full-length interview but later completed the abbreviated instrument are excluded.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

4.3.3.2 Abbreviated instrument

Abbreviated instrument overall timing. The abbreviated version of the HSLS:09 second follow-up interview included all the same sections as the full-length survey, though with fewer questions in each. On average, the abbreviated interview took 16.8 minutes to complete. Like the full-length instrument, abbreviated telephone interviews (averaging 20.7 minutes) were found to take significantly longer than abbreviated interviews that were self-administered web interviews ($t(845.28) = 17.41$, $p < .05$), which averaged 15.0 minutes. No significant timing differences were found between respondents that completed the abbreviated mobile interview (averaging 14.8 minutes) and nonmobile interview (averaging 15.2 minutes). Also sharing similarity with the full-length instrument, the timing for abbreviated field interviews (averaging 16.7 minutes) differed significantly from that for self-administered abbreviated web interviews ($t(125.61) = 3.36$, $p < .05$).

Abbreviated instrument section timing. Average section completion times were 1.2 minutes for the High School section, 4.8 minutes for Postsecondary Education, 5.7 minutes for Employment, 0.9 minutes for Family and Community, and 3.8 minutes for the Locating section.

Abbreviated self-administered web surveys took significantly less time than telephone interviews for all sections. The Postsecondary Education section was the only section that was found to be significantly longer among abbreviated web mode interviews compared to abbreviated field interviews ($t(116.41) = 2.49$, $p < .05$). Table 16 shows the average time in minutes to complete the abbreviated interview, by interview section and mode.

Table 16. Average time in minutes to complete the abbreviated interview, by interview section and mode: 2016

			Mode of administration									
	All modes		Web (nonmobile)		Web (mobile)		Web[1]		Telephone[2]		Field[3]	
Interview section	Number	Time	Number	Time	Number	Time	Number	Time	Number	Time	Number	Time
All sections[4]	**1,427**	**16.8**	**434**	**15.2**	**468**	**14.8**	**902**	**15.0**	**431**	**20.7**	**94**	**16.7**
High School	1,426	1.2	434	1.0	468	1.1	902	1.1	430	1.6	94	1.6
Postsecondary Education	1,427	4,8	434	4.9	468	4.0	902	4.4	431	5.9	94	3.6
Employment	1,426	5.7	433	5.0	468	5.2	901	5.1	431	6.8	94	6.3
Family and Community	1,426	0.9	433	0.7	468	0.7	901	0.7	431	1.3	94	1.0
Locating	1,426	3.8	433	3.1	468	3.3	901	3.2	431	5.0	94	3.6

[1] Includes self-administered nonmobile and mobile interviews.
[2] Includes centralized telephone and field telephone interviews.
[3] Includes field in-person (CAPI) and field self-administered cases.
[4] Includes timing for an additional Consent section.
NOTE: Detail may not sum to totals because of rounding. Partial interviews and cases that began the full-length interview but later completed the abbreviated instrument are excluded.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

This page intentionally left blank

Chapter 5. Data Preparation and Processing

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

5.1 Overview of Systems

Several project management software systems were used in the HSLS:09 second follow-up to develop, conduct, and support data collection in an efficient and effective manner; each system was tailored to meet the specific needs of the HSLS:09 project staff. They included an integrated management system (IMS); survey control system (SCS); computer-assisted telephone interviewing case management system (CATI-CMS); integrated field management system (IFMS); and survey development system, Hatteras. The IMS, SCS, CATI-CMS, IFMS, and Hatteras are standard RTI systems used successfully in earlier rounds of HSLS:09, as well as other National Center for Education Statistics (NCES) studies.

Integrated management system (IMS). The IMS is a comprehensive set of tools designed to provide NCES and project staff access to a centralized, easily accessible repository for project data and documents. The IMS includes tools and information used to assist project staff and NCES staff in monitoring and managing data collection. The IMS contained core information relevant to the study, including the project schedule, monthly progress reports, daily data collection reports and status reports (generated by the SCS, described below), project plans and specifications, project deliverables, instrument specifications, staff contacts, and a document archive. The IMS was accessible via the Internet and protected by Secure Sockets Layer (SSL) encryption and a password-protected login.

Survey control system (SCS). The SCS is the integrated set of databases and applications used to control and monitor all activities related to data collection, including tracing and locating of sample members. Through the SCS applications, project staff were able to perform such activities as e-mailing to groups of sample members, preparing lead letters and follow-up mailings, mail return and returned e-mail processing, executing batch tracing, reviewing locating information, tracking case statuses, and viewing comments from telephone interviewers. Applications within the SCS allowed for sample member-specific data to be used for a number of

daily tasks related to sample maintenance. Specifically, the mail-out application generated communications to sample members, the query application enabled administrators to review case-level contact information and status (in the form of U.S. Postal Service mail and e-mail), and the mail return application provided information needed to update the database as new information was received. The SCS produced various data collection monitoring reports, made available to project staff and to NCES on the IMS, which were used to track the day-to-day progress of the study's data collection overall and by subgroups of interest.

Computer-assisted telephone interviewing case management system (CATI-CMS). The CATI-CMS managed all aspects of telephone interviewing, for centralized data collection staff. The CATI-CMS included a call scheduler and a case delivery tracking system that recorded call outcomes and interviewer case notes, and a process that allowed supervisors to assign and transfer cases to interviewers. The CATI-CMS was fully integrated with the SCS so that all software systems needing sample member data (e.g., the IFMS and the IMS) accessed a common database. Case status changes in the CATI-CMS were automatically updated in the SCS during overnight processes, providing integration among all data collection systems.

Integrated field management system (IFMS). The IFMS managed all aspects of field interviewing and provided reporting processes to help supervisors track the status of field cases; see section 4.2.1 for details on the data collection design. The IFMS provided tools to distribute cases to FIs and to relay case details and completed responses to RTI. Like the CATI-CMS, the IFMS was fully integrated with the SCS so that all software systems needing sample member data accessed a common database. All IFMS data transfers were encrypted and conducted in a secure environment. A case management system (the IFMS-CMS) on field laptops enabled FIs to work cases in the field by tracking case status and event history.

Survey development system. The HSLS:09 second follow-up survey instrument was created using a web-based platform in which project staff developed, reviewed, tested, modified, and communicated changes to specifications and code for the instrument. All instrument-development specifications were stored in a structured query language (SQL) server database and were made accessible through web-browser interfaces. The survey development system provided question wording and response options, routing logic, programming, as well as testing and commenting interfaces for the HSLS:09 second follow-up survey instrument.

Each of the software systems enumerated above was equipped with safeguards to securely store and transfer personally identifiable information (PII). Processing of PII and data transfers were conducted in accordance with the Federal Information

Processing Standards (FIPS) moderate security standard. Automated processes transferred data between RTI's database located within RTI's Enhanced Security Network and the NCES database via a secure, encrypted connection. In accordance with FIPS 140.2 standards, data were encrypted prior to being transferred from RTI systems (e.g., data transfers to field laptop computers and batch tracing vendors[26]) and were decrypted once they successfully reached the destination.

5.2 Data Cleaning and Editing

The same survey instrument was used for self-administration on the Internet (via mobile and nonmobile devices) and interviewer administration (via telephone or in-person interviews). Survey response data from all administration modes were stored in an SQL database that was consistent across data collection modes. Having the same instrument database across all modes of data collection ensured that skip patterns were consistent across applications.

Project staff developed editing programs to check items for logical patterns within the questionnaire to ensure intended item dependency. For example, a respondent who indicated that he or she never attended college will not have data about his or her experiences while attending college, in which case a reserve code indicating "Item-missing, item not applicable" will represent the missing data. If a respondent indicated that he or she attended college and did not have data about his or her experiences while attending college, then those items would have a reserve code indicating "Item-missing, nonresponse." Such item dependency relationships were included in the data editing rules. Other types of edits corrected data entry errors. For example, duplicated postsecondary institution information provided by the respondent in multiple survey loops was consolidated into a single record. Additionally, relevant postsecondary institution information from Integrated Postsecondary Education Data System (IPEDS) Institution Characteristics files for 2002–15 was also added to the data file.

Some items were edited to include logical inferences, meaning the question was skipped because the answer was known based on prior responses. For example, if a student answered that she or he did not have children, a value of 0 was logically inferred for the question about the number of the student's biological children. During data processing, answers were logically inferred in these instances to make the data easier for analysts to use. Instances in which values have been logically inferred can be identified by inference flags, which take the name of the logically inferred variable and add an "_I" to the end. The SAS programming code used to

[26] For details on batch tracing vendors and data sources, see section 4.1.4.2.

construct the logical inference flags is included in the item's description window of the electronic codebook (ECB). When possible, the logic is explained using variables available in the ECB. However, in some cases the logic is based on variables containing preloaded data (i.e., data from the sample frame or a prior data collection round), temporary variables calculated in the survey, or variables from internal systems (i.e., case management system) and will not be reproducible from the data available in the ECB; item documentation will make it clear when this is the case. Table 17 provides a list of all logically inferred variables.

Table 17. Logically inferred variables: 2016

S4HIMATH	S4PRVLOANEST	S4EMPLOYER02
S4DROPOUTHS	S4PROFCERT	S4STARTJOBM2
S4TRANSFERHS	S4MLT16FB	S4STARTJOBY2
S4HSEQUEXAM	S4ACTIVEDUTY	S4OFFWORK2
S4HSEQEXAMPASS	S4ANYJOB	S4WORKNENR2
S4CHOICEAPP	S4WORKING16FB	S4LCNSE4JOB2
S4CHOICEAPPID	S4SAMEJOB1	S4UNEMPEVER
S4CHOICEACC	S4JOBENDM1	S4UNEMPDUR
S4CHOICEACCID	S4JOBENDY1	S4UNEMPFREQ
S4EVRATNDCLG	S4OFFWORK1	S4UNEMPCOMP
S4ICHOICEAPP	S4WORKENR1	S4OCC30RELATE
S4IAPPSTATUS	S4WORKHRENR1	S4BIOCHILDNUM
S4ICHOICEACC	S4WORKNENR1	S4ADPTCHILDNUM
S4ICLG16FB	S4WORKHRNENR1	S4STEPCHILDNUM
S4PPGM16FB	S4SAMEJOB2	S4INCOMECAT
S4POTHDEGMAJ	S4JOBTITLE2	S4INCOMESPCAT
S4MAJMAINRSN	S4JOBDUTY2	S4DEPCHILDNUM
S4CHGMAINRSN	S4JOB22	
S4ONCAMPUS	S4JOB62	

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

5.3 Coding

The survey instrument collected data on majors or fields of study, occupations, postsecondary institutions, and secondary schools, all of which required coding, in which text strings were classified into similar categories, or codes. The survey instrument included applications which allowed respondents or interviewers to code text strings to widely used taxonomies. To aid this process, interviewers were trained to use probing techniques to assist respondents in the coding process and web respondents were provided supporting text on screen. However, some text strings were too difficult to code during the survey or were not finalized. All text strings that were not coded during the survey were coded as part of data processing. This section describes the types of data that required coding, the applications used to code data, the coding process, quality control procedures, and measures of coding quality.

5.3.1 Major or Field of Study Coding

The survey collected the major or field of study that respondents were most seriously considering when they began postsecondary education. If the respondent reported having changed his or her major, the survey also collected the respondent's major for their reference degree (i.e., their current or most recent undergraduate degree or certificate program). To code these majors, the survey instrument included an application that allowed respondents to code their major using the NCES 2010 Classification of Instructional Programs (CIP) taxonomy.[27]

5.3.1.1 Major or field of study coding methods

To use the coding application, respondents first entered text to describe the major, after which a list of majors, customized based on the text string entered, were presented to the respondent. Respondents could then select one of the options listed or choose "none of the above." If "none of the above" was selected, a 2-tiered dropdown menu appeared. The first dropdown menu contained a general list of majors. The second dropdown menu provided more specific majors and was dependent on the selection made in the first menu. If respondents were unable to locate a suitable match, they could proceed with the survey without selecting a major.

[27] The Classification of Instructional Programs (CIP) taxonomy is organized on three levels: (1) the 2-digit series, (2) the 4-digit series, and (3) the 6-digit series. The 2-digit series represents the most general groupings of related programs. The 4-digit series represents intermediate groupings of programs that have comparable content and objectives. The 6-digit series, also referred to as 6-digit CIP codes, represents specific instructional programs. For example, "International Economics" has the 6-digit code 45.0605, which places it in "Economics" (4-digit CIP 45.06) and "Social Sciences" (2-digit CIP 45). Only the 2-digit and 6-digit CIP code levels were used for instrument coding analysis and quality control within the second follow-up. For more details on the CIP 2010 taxonomy, see: https://nces.ed.gov/ipeds/cipcode/Default.aspx?y=55.

In this case, the text string and any selections from the dropdown menus were retained.

All major text strings that were not coded in the survey instrument were processed by survey staff. First, the major text strings that were reported more than once (i.e., by multiple respondents or multiple times for a single respondent) were assigned a code by an expert coder (EC). This code was then applied to all other exactly matching text strings, or *batch coded*, to ensure consistency of codes for duplicated text strings. Of the 2,862 major text strings to be coded, 36 percent were batch coded. The remaining text strings were then *upcoded* to the CIP taxonomy by ECs using an application that used the same search function as the coding application in the instrument. The EC could assign a CIP code or indicate that the text string was too vague to code. Additionally, if a major was coded only at the 2-digit level (e.g., 45), then a 6-digit code was assigned (e.g., 45.000). Table 18 provides the survey-uncoded majors, by coding method.

Table 18. Survey-uncoded majors, by coding method: 2016

	Total text strings		Batch coded		Coded by expert coders	
Coding system	Number	Percent	Number	Percent	Number	Percent
Survey-uncoded majors	2,862	100.0	1,030	36.0	1,832	64.0

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

5.3.1.2 Major or field of study coding quality control procedures and results

To evaluate the quality of the coding completed during the survey, a random sample of approximately 10 percent of majors coded during the survey were selected for recoding and analysis. In the recoding procedure, survey staff evaluated text strings and assigned CIP codes without knowledge of the codes selected during the survey. If the EC selected a code that differed from that selected during the survey, the EC was then shown both codes. The EC was instructed to recode using the different code only if the code selected during the survey was clearly incorrect. When the EC disagreed with the CIP code selected during the survey, that case was sent to a second EC for adjudication. If a code selected in the survey 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. After adjudication, 101 major text strings (4 percent) were designated as "too vague to code."

Approximately 95 percent of the codes selected during the survey were deemed to be accurate at the most detailed 6-digit CIP code level. The EC disagreed at the 2-digit

level with the CIP code selected during the survey for only 2 percent of the strings. Results of the major recoding process are given in table 19.

Table 19. Results of quality control recoding and upcoding of major: 2016

Results	Number	Percent
Sample of strings coded during survey	1,638	100.0
Match at 6-digit and 2-digit	1,551	94.7
Match at 2-digit, but not 6-digit	62	3.8
Disagree	25	1.5
Sample of strings coded during data processing	194	100.0
Match at 6-digit and 2-digit	62	32.0
Match at 2-digit, but not 6-digit (coded to 6 digits)	25	12.9
Match at 2-digit, but not 6-digit (coded to 2 digits)	10	5.2
Match at too vague to code	14	7.2
Disagree	83	42.8

NOTE: Detail may not sum to totals because of rounding. After adjudication, a total of 101 major text strings were classified as "too vague to code."
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Upcoded strings were also subject to a quality control review and analysis. All upcoded strings were selected for independent coding by a second EC and the results of the two ECs were compared. Results of this upcoding review are also provided in table 19. The two ECs selected the same detailed 6-digit code for 32 percent of the upcoded text strings. For 13 percent of the text strings, the ECs agreed at the 2-digit level but disagreed at the detailed 6-digit level. There were also 10 additional cases (5 percent) for which the two ECs agreed, though the cases were only coded to the 2-digit level due to difficulty with locating an accurate 6-digit code. Additionally, both ECs determined that the text was too vague to code for 7 percent of the upcoded strings. Disagreement between the two coders occurred for 43 percent of the upcoded strings. It should be noted that the text strings that were not coded during the survey were the most difficult cases to code. All instances in which there was disagreement at either the 6-digit or 2-digit level, or where one EC identified the string as too vague to code, were adjudicated by a third EC. If the adjudicating EC disagreed with the both other ECs, the final code determination was made by the adjudicator.

5.3.2 Occupation Coding

The second follow-up instrument included a tool that allowed occupation job titles and duties to be matched to occupation descriptions from the U.S. Department of Labor, Employment & Training Administration's Occupational Information

Network (O*NET) which uses the 2010 Standard Occupational Classification (SOC) taxonomy.[28]

5.3.2.1 Occupation coding methods

Respondents were asked to provide a job title and job duties for each occupation reported. These text strings were then automatically matched to the occupation descriptions from O*NET and a customized list of occupations was presented. 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 respondent with a set of three sequential dropdown menus, each with choices increasing in their level of specificity dependent on the previous selection. If the respondents were unable to find an appropriate O*NET match for the occupation, they could proceed with the survey without selecting a job title. In this case, the text string and any selections from the dropdown menus were retained to assist with coding during data processing.

ECs attempted to code all occupations that were not coded in the survey. This upcoding was completed using an application that used the same search features as the application in the survey instrument. The EC could assign an O*NET code or indicate that the text string was too vague to code. Additionally, if an occupation was coded only at the 2-digit level (e.g., 12), then a 6-digit code was assigned (e.g., 12.000).

5.3.2.2 Occupation coding quality control procedures and results

ECs evaluated the quality of coding that was completed during the survey by recoding a random sample of approximately 10 percent of the occupations. To recode the sampled occupations, staff members worked with a coding application which used the same search function as the application in the instrument. ECs evaluated text strings and assigned codes without knowledge of the codes that were selected during the survey. If the code selected differed from the code assigned during the survey, the EC was then shown both codes. The EC was instructed to only override the code selected in the survey if it was clearly incorrect. When the EC did not agree with the 6-digit code selected during the survey, the string was adjudicated by a second EC. Text strings were designated "too vague to code" when

[28] O*NET employs a taxonomy that provides a nested coding structure: 23 two-digit codes expand to 96 four-digit codes that can be expanded further to 821 six-digit code categories. For more details, see: https://www.onetonline.org/.

they lacked sufficient clarity or specificity. After adjudication, 81 occupation text strings (2 percent) were designated as too vague to code.

ECs agreed with the 6-digit code selected during the survey for 94 percent of the text strings reviewed and agreed with the 2-digit code (but not the 6-digit code) for an additional 3 percent of the text strings reviewed. Note that if an occupation was coded at only the 2-digit level (e.g., 26) in the interview, then a 6-digit code was assigned (e.g., 26.0000). ECs disagreed with the 2-digit code selected during the survey for 3 percent of the occupations. The results of occupation recoding are displayed in table 20.

Table 20. Results of quality control recoding and upcoding of occupation: 2016

Results	Number	Percent
Sample of strings coded during survey	1,942	100.0
Match at 6-digit and 2-digit	1,820	93.7
Match at 2-digit, but not 6-digit	55	2.8
Disagree	67	3.5
Sample of strings coded during data processing	2,346	100.0
Match at 6-digit and 2-digit	1,159	49.4
Match at 2-digit, but not 6-digit (coded to 6 digits)	435	18.5
Match at 2-digit, but not 6-digit (coded to 2 digits)	36	1.5
Match at too vague to code	35	1.5
Disagree	681	29.0

NOTE: Detail may not sum to totals because of rounding. After adjudication, a total of 81 occupation text strings were classified as "too vague to code."
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

All the upcoded occupation strings were also selected for independent coding by a second EC. The results of this independent coding are also shown in table 20. The two ECs selected the same detailed 6-digit code for 49 percent of the text strings and selected the same 2-digit code (but not the 6-digit code) for 19 percent of the upcoded strings. There were 36 cases (2 percent) in which the ECs agreed with one another for strings that were only coded to the 2-digit level because an accurate match could not be found at the 6-digit level. A small percentage of upcoded strings were deemed too vague to code by both ECs (1 percent). The coders disagreed for 29 percent of the occupations. Note that the occupations that were not coded during the survey are the most difficult strings to code. Cases in which ECs disagreed at either the 6-digit or 2-digit level, or one EC identified the occupation as too vague to code, were adjudicated by a third EC. If the adjudicating EC disagreed with the other ECs' code selections, the final code determination was made by the adjudicator.

5.3.3 Respondent Job at Age 30 Coding

The second follow-up survey instrument asked respondents to indicate what occupation they thought they would have when they turned 30 years old. Respondents entered a job title in the survey but were not asked to report expected job duties due to the hypothetical nature of the job. Respondents were also provided the option to select a job title previously reported in the survey, or to indicate that they did not know what job they would hold or did not plan on working at age 30.

5.3.3.1 Respondent job at age 30 coding methods

Respondents were not asked to code their expected occupations in the survey, therefore all job titles were coded by survey staff after data collection using the O*NET taxonomy. Duplicated text strings were first batch coded; of the 10,087 text strings provided by respondents, 9,340 strings (93 percent) were batch coded. The 747 text strings that remained uncoded after batch coding were coded by ECs using the O*NET coding application described in 5.3.2.1. However, because job duties were not collected for respondents' expected jobs at age 30, the search for an appropriate O*NET code was strictly based on the job title. Table 21 displays the results of the expected job at age 30, by coding method.

Table 21. Job at age 30 results, by coding method: 2016

	Total text strings		Batch coded		Coded by expert coders	
Coding system	Number	Percent	Number	Percent	Number	Percent
Job at age 30 text strings	10,087	100.0	9,340	92.6	747	7.4

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

5.3.3.2 Respondent job at age 30 coding quality control procedures and results

Because the survey instrument did not have respondents code the job title for the job at age 30, all such text strings were coded by two ECs. The two ECs arrived at the same results (a match at the 6- and 2-digit levels) for 32 percent of the text strings and agreed on the same 2-digit code (but not the 6-digit code) for 15 percent. Both coders determined the string was too vague to code for 11 percent of the cases. Disagreement occurred at the 2-digit level in 42 percent of the cases. This higher percentage (relative to the occupation code disagreement rate reported in section 5.3.2.2) was expected given that respondents were not asked to provide job duties and therefore the ECs had less information at their disposal. When ECs disagreed (either at the 6- or 2-digit level), cases were sent to a third EC for adjudication. If the adjudicator disagreed with the both of the other ECs, the final code determination

was made by the adjudicator. Table 22 displays results of the job at age 30 upcoding process.

Table 22. Results of quality control of job at age 30 upcoding: 2016

Results	Number	Percent
Total coded by second expert coder	747	100.0
Match at 6-digit and 2-digit	238	31.9
Match at 2-digit, but not 6-digit	111	14.9
Match at too vague to code	82	11.0
Disagree	316	42.3

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

5.3.4 High School Coding

The second follow-up survey collected the name, city, and state of high schools respondents had attended, other than high schools reported in previous rounds of data collection. The survey requested that respondents provide the name of the high school they last attended if it was different than one reported in prior rounds of HSLS:09. Respondents were asked to code high schools during the survey using a coding application embedded within the survey. The coding application matched high school information 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 included in the application to account for schools that may have closed or opened over the span of time between the HSLS:09 base-year and second follow-up surveys. Any high schools left uncoded in the survey were later coded by ECs during subsequent data processing. In total, 507 high schools were provided in second follow-up surveys. Of these, 336 were coded during the survey and 171 were left uncoded in the survey.

Note that survey staff did not attempt to recode survey-coded schools due to the greater objectivity of coding high schools, compared to coding majors or occupations. When coding high schools, the respondent was able to see the high school name, city, and state; respondents are generally familiar with the name of the school they attended, whereas majors or occupation titles can be described differently depending on the school or employer, making these harder to match exactly in the coder database.

5.3.4.1 High school coding methods

Coding personnel used the high school coding application to code survey-uncoded high schools. ECs also conducted Internet searches to locate schools that were not found in CCD or PSS; these tended to be alternative schools, charter schools, or high schools that had recently opened.

5.3.4.2 High school coding results

High school coding results are presented in table 23. Fewer than half of the survey-uncoded high schools, 44 percent, were found to be in the United States.

Table 23. Final disposition of survey-uncoded high schools, after coding: 2016

Results	Number	Percent
Survey-uncoded high schools	171	100.0
Located in United States	76	44.4
Unlocated or foreign	95	55.6

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

5.3.5 Postsecondary Institution Coding

In the second follow-up, respondents were asked to indicate any postsecondary institutions that they had attended after high school through February 2016 (the survey reference date), up to two other institutions to which they had applied, any institutions the respondent first attended after February 2016, and any institutions the respondent planned to attend between the survey date and December 2016. Institutions not attended by the end of February 2016 were coded but not included in the data file because attendance began after the survey reference date. However, postsecondary transcripts were collected for these institutions in 2017. Postsecondary institutions provided by respondents were coded during the survey using the IPEDS universe of institutions.[29] In total, 13,623 postsecondary institutions were provided in the second follow-up. Of these, 12,328 were coded during the survey and 1,295 were left uncoded in the survey.

As with high school coding, survey staff did not attempt to recode survey-coded postsecondary institutions due to the greater objectivity of coding postsecondary institutions, compared to coding majors or occupations. When coding postsecondary institutions, the respondent was able to see the postsecondary institution name, city,

[29] IPEDS is a system of interrelated surveys conducted annually that gather information from all postsecondary institutions that participate in Title IV federal student financial aid programs. For more details, see: https://nces.ed.gov/ipeds/.

and state; respondents are generally familiar with name of the institution they attended.

5.3.5.1 Postsecondary institution coding methods

After respondents entered the institution's name, city, and state into the survey, they could use a coding application containing institution data from IPEDS to identify and code the appropriate institution. When a match was not found, the respondent was asked to provide the institution's level (i.e., 4-year, 2-year, or less-than-2-year) and control (i.e., public, private not-for-profit, or private-for-profit). This information was later used to assist coding staff in identifying a match in IPEDS as part of data processing.

Only text strings not coded during the survey were processed for upcoding. Cases with an institution name, city, and state that exactly matched an IPEDS record, but had not been coded during the survey, were automatically assigned the corresponding IPEDS ID, or auto-coded; 13 percent of the survey-uncoded institutions were able to be auto-coded. Remaining uncoded cases were loaded into the coding application for an EC to assign IPEDS IDs. Table 24 displays survey-uncoded postsecondary institutions by coding method.

Table 24. Survey-uncoded postsecondary institutions, by coding method: 2016

	Total text strings	Auto-coded		Coded by expert coders	
	Number	Number	Percent	Number	Percent
Survey-uncoded postsecondary institutions	1,295	168	13.0	1,127	87.0

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

5.3.5.2 Postsecondary institution coding results

Table 25 provides the final disposition of the survey-uncoded postsecondary institutions. After auto-coding and EC coding, 62 percent of the institutions were found to be located in the United States and were successfully assigned an IPEDS ID, 20 percent were unlocated, 7 percent identified as foreign institutions, and 11 percent were identified as not postsecondary institutions.

Table 25. Final disposition of survey-uncoded postsecondary institutions, after coding: 2016

Results	Number	Percent
Survey-uncoded postsecondary institutions	1,295	100.0
Located in United States	797	61.5
Unlocated	263	20.3
Foreign	92	7.1
Not a postsecondary institution	143	11.0

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Chapter 6. Response Rates, Analytic Weights, Variance and Design Effects Estimation, Nonresponse Bias Analysis, Imputation, and Disclosure Avoidance

The post-data collection statistical activities conducted to support the analysis of second follow-up data are presented and discussed in this chapter. A discussion of weighted unit response rates from each round of the High School Longitudinal Study of 2009 (HSLS:09) is provided in section 6.1. A succinct description of the weights developed prior to the second follow-up is provided in section 6.2, an in-depth discussion of the weights developed for the second follow-up is provided in section 6.3, and guidance on the process of selecting weights for particular analyses is provided in section 6.4. The appropriate calculation of standard errors and estimates of the impact of sampling and weight adjustments on the precision of standard errors is discussed in section 6.5. Item-level declined response is presented in section 6.6, and a discussion of bias arising from item nonresponse and unit nonresponse is given in section 6.7. The methods and results of imputation procedures are presented in section 6.8. Section 6.9 discusses the application of disclosure limitation techniques and explains the resulting differences between public-use and restricted-use data files.

6.1 Unit Response Rates

Information on the participation of HSLS:09 sample members is of interest to understand the data collection effort and data quality. As discussed in section 4.3, in calculating response rates there are two types to consider: *participation rate* and *response rate*. Section 4.3 provides coverage of participation rates, and response rates are detailed in this section.

For the HSLS:09 study, weighted unit response rates computed using the base weights are used to gauge the degree to which participating schools and participating students represent their respective populations. When response rates are higher, the collected data may produce more accurate population estimates, because the larger responding sample is likely to better represent the target population of interest. The

weighted unit response rates reported in this data file documentation are calculated using the response rate formula provided in National Center for Education Statistics (NCES) Statistical Standard 1-3-2 (Seastrom 2014).

Calculation of a weighted response rate requires identifying the population of interest (school or student) and specifying a participation definition. In studies such as HSLS:09 that are longitudinal in nature and utilize multiple survey components in one or more study round, there are a multitude of participation definitions that may be created. For example, a student participant may be defined as a student who completed the HSLS:09 base-year student questionnaire and responded to the second follow-up survey or, alternatively, a student participant may simply be defined as a student who completed the base-year student questionnaire. Several weighted unit response rates, using different definitions of participation and covering all HSLS:09 study rounds, are provided in this section.

Although higher response rates can indicate more accurate survey results, it is also important to examine whether there is the potential for nonresponse bias to exist in the data. NCES standards require unit nonresponse bias analyses to be conducted when weighted unit response rates fall below 85 percent. Use of analytic weights enables population estimates to be calculated from sample data. The base weights adjust for differential selection probabilities.

For some of the survey components in each of the HSLS:09 study rounds, weighted unit response rates computed using the base weights are provided in table 26 as an overview; for a complete listing, see table 30 in section 6.4. Note that schools and students are the sampling units, not parents; accordingly, response rates are interpreted with respect to schools and students. Weighted response rates incorporating base-year teacher data and rates incorporating multiple sets of data across more than one study round are provided in section 6.4.

Table 26. HSLS:09 Base-weighted Unit Response Rates

Unit	Participation definition	Eligible	Participated	Weighted percent
	Base Year			
School	School agreed to participate	1,889	944	55.5 [1]
Student	Student questionnaire completed	25,206	21,444	85.7 [2]
	Student assessment completed	25,206	20,781	83.0 [2]
	First Follow-up			
Student	Student questionnaire completed [3]	25,184	20,594	82.0 [2]
	Student assessment completed [3]	25,184	18,507	73.0 [2]
	Parent questionnaire completed [5]	11,952	8,651	72.5 [4]
	2013 Update and High School Transcript components			
Student	Student questionnaire completed	25,168	18,558	73.1 [2]
	High school transcripts collected	25,167	21,928	87.7 [2]
	Student questionnaire completed and high school transcripts collected	25,167	17,656	70.2 [2]
	Second Follow-up			
Student	Student questionnaire completed	25,123	17,335	67.9 [2]

[1] Weighted percentage is calculated using the school base-weight.
[2] Weighted percentages are calculated using the student base-weight.
[3] A total of 22 students from the base year were ineligible for the first follow-up.
[4] Weighted percentage is calculated using the student base-weight adjustment for parent subsampling.
[5] A subsample of 11,952 eligible parents were asked to participate in the HSLS:09 first follow-up data collection.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

As shown in table 26, weighted response rates for the student questionnaire, which is the only component included in all four data collections, ranged from a high of 85.7 percent in the base year to a low of 67.9 percent in the second follow-up.

6.2 Overview of Weighting in the Base Year, First Follow-up, and 2013 Update

The use of weights is essential to produce estimates that are representative of the HSLS:09 target population of students. An analysis weight should be used to produce survey estimates. When testing hypotheses (e.g., conducting *t* tests and regression analyses) using weighted data from a study such as HSLS:09 that has a complex design, analysts also should use methods to properly estimate variances. Variables have been created for HSLS:09 to support two methods of variance estimation that account for the HSLS:09 complex sample design: (1) a balanced repeated replication (BRR) variance estimation method using the BRR weights and the associated analytic weight and (2) a linearization variance estimation method through a Taylor series approximation using analytic weights and variables that represent school sampling strata and primary sampling units. For more details on standard error estimation, see section 6.5.

Five sets of weights were constructed for the HSLS:09 base year: a school-level weight to analyze information collected in the administrator and counselor questionnaires as well as school-level data from other sources, such as the Common Core of Data (CCD) and Private School Universe Survey (PSS), linked to participating schools; a student-level weight to analyze student survey responses and mathematics assessment scores; and three contextual weights to analyze responses obtained from the science teacher questionnaire, the mathematics teacher questionnaire, and the home-life (parent) questionnaire. The steps implemented to create these weights are detailed in the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011).

Four analytic weights were computed for the HSLS:09 first follow-up using methods similar to those implemented in the base year. Two of the four are student-level weights: one for analyses specific to the first follow-up student survey and one for longitudinal analyses examining both base-year and first follow-up student survey data. The other two are home-life contextual weights: one for analyses of only data from the first follow-up parent questionnaire and one for longitudinal analyses of data from both the base-year and first follow-up parent questionnaires. The steps utilized to create these weights are discussed in detail in the *HSLS:09 Base Year to First Follow-Up Data File Documentation* (Ingels et al. 2013).

Nine analytic weights—four nontranscript weights and five transcript weights—were computed for the HSLS:09 2013 Update and High School Transcript study using methods similar to those implemented in the base year and first follow-up. The four nontranscript weights are designed for the following analyses:

- analyses specific to the 2013 Update;
- analyses examining both base-year and 2013 Update data;
- analyses examining both the first follow-up and the 2013 Update; and
- analyses examining base-year, first follow-up, and the 2013 Update data.

The five transcript weights are designed for the following analyses:

- analyses specific to High School Transcript data only;
- analyses that utilize 2013 Update data combined with High School Transcript data;
- analyses that incorporate base-year, 2013 Update, and High School Transcript data;
- analyses that examine first follow-up, 2013 Update, and High School Transcript data; and
- analyses that use base-year, first follow-up, 2013 Update, and High School Transcript data.

The steps used to construct these weights are detailed in the *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015).

6.3 Second Follow-up and Supplemental 2013 Update Weights

Seven analytic weights were computed for data from the HSLS:09 second follow-up using a similar methodology as implemented in the base-year, first follow-up, 2013 Update, and High School Transcript collections. Five of the weights were constructed for analyzing data from the second follow-up student interview survey, and two weights were constructed for analyzing data from the 2013 Update. The two weights for 2013 Update respondents account for base-year and 2013 Update student survey response in conjunction with base-year teacher response and supplement existing 2013 Update weights. The second follow-up weights are discussed in detail within section 6.3.1, the supplemental 2013 Update weights are detailed in section 6.3.2, and quality control measures employed to construct these weights are discussed in section 6.3.3.

6.3.1 Second Follow-up Weights

An overview of the five second follow-up weights, which provides information about the analyses for which each weight is appropriate, is included in section 6.3.1.1. Section 6.3.1.2 includes a high-level discussion of the computation of the second follow-up weights. More detail regarding the computation of the second follow-up weights can be found in appendix H. Characteristics of the five second follow-up weights, including summary statistics, can be found in section 6.3.1.3.

6.3.1.1 Overview of second follow-up weights

The five analysis weights for analyzing data from the second follow-up include one for analyses specific to the second follow-up (W4STUDENT); one for analyses using data from the base-year and second follow-up student surveys (W4W1STU); one for analyses using data from the base-year, first follow-up, 2013 Update, and second follow-up student surveys (W4W1W2W3STU); one for analyses using data from the base-year and second follow-up student surveys as well as the base-year parent questionnaire (W4W1STUP1); and one for analyses using data from the base-year and second follow-up student surveys and the base-year and first follow-up parent questionnaires (W4W1STUP1P2).

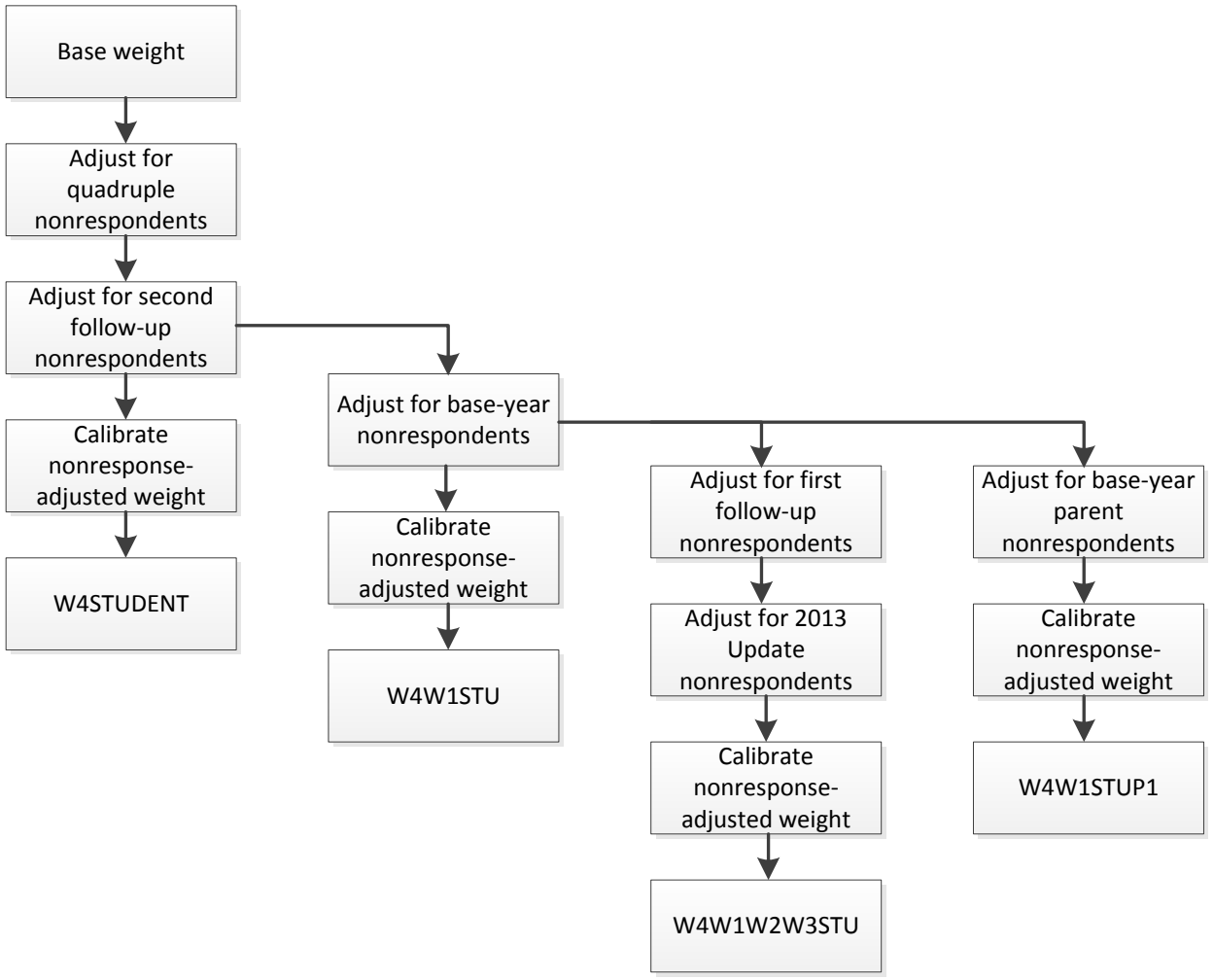
6.3.1.2 Computation of second follow-up survey weights

Two types of weight adjustments are used to produce the five analysis weights and associated BRR weights for the second follow-up. The construction of the analysis weights is discussed in section 6.3.1.2.1 and the construction of the corresponding BRR weights is discussed in section 6.3.1.2.2.

6.3.1.2.1 Analysis weights

Four of the five second follow-up analysis weights (W4STUDENT, W4W1STU, W4W1W2W3STU, and W4W1STUP1) were constructed in a manner such that the adjustments for nonresponse were performed in a sequential fashion with certain adjustments applied to multiple weights. That is, one weight adjustment may have been used in the construction of multiple weights. For example, the adjustment for quadruple nonresponse, adjusting for students who were nonrespondents in the base year, first follow-up, 2013 Update, and second follow-up, was used in the construction of W4STUDENT, W4W1STU, W4W1W2W3STU, and W4W1STUP1. Figure 9 displays a visual representation of these sequential weighting adjustments.

Figure 9. Second follow-up weight construction process for the four weights with sequential adjustments for nonresponse



NOTE: Unless otherwise specified, adjustments for nonresponse are based on student survey nonresponse. The nonresponse adjustments occurring after the adjustment for quadruple nonrespondents are designed to account for nonresponse in the stated round among sample members who responded in one or more prior rounds.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

In the construction of the four weights with sequential adjustments, two nonresponse adjustments were used:

A1. adjustment to the student base weight[30] for student questionnaire nonresponse to all four rounds of HSLS:09; and

[30] Base weights compensate for unequal probabilities of selection into the study sample. A base weight is calculated as the inverse probability of selection and includes all stages of sample design (e.g., two design stages are used for HSLS:09). For more details on base weights, see appendix A.

A2. adjustment to the weight in (A1) for remaining student questionnaire nonresponse to the second follow-up survey.

To finalize the second follow-up student analytic weight (W4STUDENT), the weight adjusted for student questionnaire nonresponse in the second follow-up was then calibrated to the same student control totals defined in the base year and used in the weight calibration for the base-year, first follow-up, 2013 Update, and High School Transcript collections.

An additional nonresponse adjustment was used in the construction of three analytic weights, W4W1STU, W4W1W2W3STU, and W4W1STUP1:

A3. adjustment to the weight in (A2) for student questionnaire nonresponse in the base year.

To finalize the base year to second follow-up student analytic weight (W4W1STU), the weight adjusted for quadruple student questionnaire nonresponse, student questionnaire nonresponse in the second follow-up, and student questionnaire nonresponse in the base year, was then calibrated to the same student control totals used in all prior rounds of the study.

For the construction of the base year to first follow-up to 2013 Update to second follow-up student analytic weight (W4W1W2W3STU), two additional nonresponse adjustments were performed:

A4. adjustment to the weight in (A3) for student questionnaire nonresponse in the first follow-up; and

A5. adjustment to the weight in (A4) for questionnaire nonresponse in the 2013 Update.

To finalize W4W1W2W3STU, the weight adjusted for quadruple questionnaire nonresponse, student questionnaire nonresponse in the second follow-up, student questionnaire nonresponse in the base year, student questionnaire nonresponse in the first follow-up, and questionnaire nonresponse in the 2013 Update was then calibrated to the same student control totals used in all prior rounds of the study.

Data from parents of sampled students were collected in the base year and data from a parent subsample were collected in the first follow-up. The 2013 Update involved collecting data from either the student or the parent. There was no parent data collection in the second follow-up. Two second follow-up weights were constructed to incorporate parent nonresponse in the base year and first follow-up when

analyzing parent data with second follow-up data. W4W1STUP1 accounts for student nonresponse in the base year and first follow-up and parent nonresponse in the base year. W4W1STUP1P2 accounts for both student and parent nonresponse in the base year and first follow-up.

For the construction of W4W1STUP1, one additional nonresponse adjustment was performed on the weight adjusted for quadruple student questionnaire nonresponse, student questionnaire second follow-up nonresponse, and student questionnaire nonresponse in the base year:

> A6. adjustment to the weight in (A3) for parent questionnaire nonresponse in the base year.

To finalize the base year to second follow-up student analytic weight that accounted for base-year parent questionnaire nonresponse (W4W1STUP1), the weight adjusted for quadruple student questionnaire nonresponse, student questionnaire nonresponse in the second follow-up, student questionnaire nonresponse in the base year, and parent questionnaire nonresponse in the base year was then calibrated to the same student control totals used in all prior rounds of the study.

Due to the subsampling used to restrict the set of students for whom parent surveys were pursued in the first follow-up, the fifth weight—the base year to second follow-up analytic weight that accounted for base-year and first follow-up parent response (W4W1STUP1P2)—was constructed independently from the other four second follow-up weights.[31] The adjustments performed to construct W4W1STUP1P2 were

> B1. The student base weight was adjusted for subsampling of parents in the first follow-up.
>
> B2. The weight in (B1) was adjusted for student questionnaire nonresponse in the base year.
>
> B3. The weight in (B2) was adjusted for student questionnaire nonresponse in the second follow-up.
>
> B4. The weight in (B3) was adjusted for parent questionnaire nonresponse to the base-year parent questionnaire, the first follow-up questionnaire, or both.

To finalize the base year to second follow-up analytic weight that accounted for base-year and first follow-up parent nonresponse (W4W1STUP1P2), the weight

[31] For more details on the subsampling of parents in the first follow-up, please refer to the *HSLS:09 Base Year to First Follow-Up Data File Documentation* (Ingels et al. 2013).

adjusted for parent subsampling in the first follow-up, student questionnaire nonresponse in the base year, student questionnaire nonresponse in the second follow-up, and parent questionnaire nonresponse in either the base year or first follow-up or both was calibrated to the same student control totals used in all prior rounds of the study. The adjustment for subsampling of parents in the first follow-up was done first to account for student eligibility before accounting for nonresponse.

For all five second follow-up weights, unit nonresponse adjustments incorporated student-level and school-level characteristics where possible using the WTADJUST procedure in SUDAAN. The calibrations for each weight also used the WTADJUST procedure in SUDAAN.

For more detail on the construction of all five second follow-up weights, please refer to appendix H. Additional information on using the analysis weights to estimate standard errors is provided in section 6.5.1.

6.3.1.2.2 Balanced repeated replication weights

A set of 200 BRR weights was created for each of the five second follow-up analytic weights. These sets of BRR weights include (1) second follow-up student weights (W4STUDENT001–200); (2) base-year to second follow-up student weights (W4W1STU001–200); (3) base-year to first follow-up student weights, to 2013 Update to second follow-up student weights (W4W1W2W3STU001–200); (4) base-year to second follow-up student weights with an adjustment for base-year parent nonresponse (W4W1STUP1001–200); and (5) base-year to second follow-up student weights with adjustments for base-year and first follow-up parent nonresponses (W4W1STUP1P2001–200). Procedures for constructing the weights mirrored those used to construct the corresponding analytic weight. Namely, the BRR weights were constructed by subjecting the base-year BRR base weights, defined for each of 200 replicates, to nonresponse and calibration adjustments following a process like that used to develop the analysis weights. Additional information on using the BRR weights to estimate standard errors may be found in section 6.5.1.

6.3.1.3 Characteristics of second follow-up survey weights

The characteristics of the five second follow-up analytic weights are presented in table 27. For each weight, the number of respondents, the average weight, the standard deviation, the minimum and maximum, and weight sums are provided.

Table 27. Descriptive characteristics of second follow-up survey weights

Weight	Number of respondents	Mean	Standard deviation	Minimum	Maximum	Sum[1]
W4STUDENT	17,335	241.3	311.01	2.0	7,891.9	4,183,284
W4W1STU	15,909	259.8	343.15	2.0	7,948.9	4,133,580
W4W1W2W3STU	13,283	311.2	412.39	2.6	9,243.2	4,133,878
W4W1STUP1	12,888	322.6	427.37	3.1	10,131.5	4,157,769
W4W1STUP1P2[2]	5,427	765.3	976.31	6.5	18,329.0	4,153,488

[1] The student counts in table 10 of chapter 3 in the *HSLS:09 Base Year Data File Documentation* (Ingels et al. 2011) were used as the control totals. Weight sums differ from the population counts because of the suppression of data from the public-use file for students who were excluded from the base-year or first follow-up student survey because it was not offered in a format that allowed their meaningful participation (students referred to as "questionnaire incapable" in response status variables) in the base year or first follow-up and deceased students being included in the calibration and subsequently having their weights set to zero.
[2] Respondents for this weight are restricted to a subset of those students whose parents were selected for the parent subsample in the first follow-up. For more detail on the first follow-up parent subsample, please refer to section 3.3.4 of the *HSLS:09 Base Year to First Follow-up Data File Documentation* (Ingels et al. 2013).
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up, Public-use Data File.

6.3.2 Supplemental 2013 Update Weights

Two supplemental 2013 Update weights were created to be used in analysis of 2013 Update data in conjunction with base-year student responses and base-year math and science teacher responses. These weights may be used, for example, to examine student outcomes at the modal high school completion time point while controlling for base-year student characteristics.

The two supplemental 2013 Update weights include a weight for analyses using base-year student survey data, 2013 Update data, and base-year math teacher data (W3W1MATHTCH); and one for analyses using base-year student survey data, 2013 Update survey data, and base-year science teacher questionnaire data (W3W1SCITCH).

Section 6.3.2.1 includes a high-level discussion of the computation of the supplemental 2013 Update weights. More detail regarding the computation of the supplemental 2013 Update weights can be found in appendix H. Characteristics of the two supplemental 2013 Update weights, including summary statistics, can be found in section 6.3.2.2.

6.3.2.1 Computation of supplemental 2013 Update survey weights

This section discusses the two types of weight adjustments used to produce the two analysis weights and associated BRR supplemental 2013 Update weights.

6.3.2.1.1 Analysis weights

Construction of the two supplemental 2013 Update analysis weights followed a process like that used to construct the base-year weights W1SCITCH and W1MATHTCH[32] in that a single weight adjustment to calibrate weight sums to control totals defined in the base year was applied to an existing student analysis weight (W3W1STU[33] in the case of these supplemental 2013 Update weights).

Specifically, for the first supplemental base year to 2013 Update weight (W3W1MATHTCH), a single weight adjustment was applied to the subset of students with a nonzero value of W3W1STU who were either not enrolled in a math course in the base year or who were enrolled in a math course in the base year and for whom a math teacher response was collected. The single weight adjustment applied to W3W1STU for this subgroup of students was designed to ensure that the weight sums of the adjusted weight matched the weight sums of W3W1STU. Students who were not enrolled in a math course in the fall of the base year subsequently had their weights set to zero.

For the second supplemental base year to 2013 Update weight (W3W1SCITCH), a single weight adjustment was applied to the subset of students with a nonzero value of W3W1STU who were either not enrolled in a science course in the base year or who were enrolled in a science course in the base year and for whom a science teacher response was collected. The single weight adjustment applied to W3W1STU for this subgroup of students was designed to ensure that the weight sums of the adjusted weight matched the weight sums of W3W1STU. Students who were not enrolled in a science course in the fall of the base year subsequently had their weights set to zero.

For both supplemental 2013 Update weights, calibrations incorporated student-level and school-level characteristics where possible using the WTADJUST procedure in SUDAAN. For more detail on the calibration adjustments used in the construction of the two supplemental 2013 Update weights, please refer to appendix H.

---

[32] For more detail on the construction of W1SCITCH and W1MATHTCH, please refer to the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011).

[33] W3W1STU accounts for (1) base-year school nonresponse and (2) student nonresponse in both the base year and the 2013 Update. All records for sample students who participated in the base year and 2013 Update will have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students adjusted for the number of deceased students observed in the HSLS:09 sample. For more detail on W3W1STU, please refer to the *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015).

6.3.2.1.2 Balanced repeated replication weights

A set of 200 BRR weights was created for each of the supplemental 2013 Update analytic weights. These sets of BRR weights include (1) base year to second follow-up with base-year math teacher weights (W3W1MATHTCH001–200) and (2) base year to second follow-up with base-year science teacher weights (W3W1SCITCH001–200). Procedures for constructing the weights mirrored those used to construct the corresponding analytic weight. Namely, BRR base weights were constructed and subjected to calibration adjustments developed for each replicate.

6.3.2.2 Characteristics of supplemental 2013 Update survey weights

The characteristics of the two supplemental 2013 Update analytic weights are presented in table 28. For each weight, the number of respondents, the average weight, the standard deviation, the minimum and maximum, and weight sums are provided.

Table 28. Descriptive characteristics of supplemental 2013 Update survey weights

Weight	Number of respondents	Mean	Standard deviation	Minimum	Maximum	Sum[1]
W3W1MATHTCH	12,812	315.5	429.33	3.5	10,322.7	4,042,751
W3W1SCITCH	11,803	319.0	440.15	2.4	10,900.4	3,764,708

[1] The student counts in table 10 of chapter 3 in the *HSLS:09 Base Year Data File Documentation* (Ingels et al. 2011) were used as the control totals. Weight sums differ from the population counts because of the suppression of data from the public-use file for students who were excluded from the base-year student survey because it was not offered in a format that allowed their meaningful participation (students referred to as "questionnaire incapable" in response status variables) in the base year and deceased students being included in the calibration and subsequently having their weights set to zero. Values may not sum to overall totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up, Public-use Data File.

6.3.3 Weighting Quality Control

A good weight is one which allows an analyst to adequately adjust for unit nonresponse and adjust for frame coverage through calibration while minimizing overall weight variability. To assess the quality of the analysis weights and their corresponding sets of replicate weights, staff reviewed the following:

- the initial base weights' characteristics, including the (1) distribution of the weights, (2) ratios of maximum weights to minimum weights, (3) unequal weighting effects, (4) ranges of weight adjustment factors, and (5) weight sums;
- the weight adjustment factors used to produce the second follow-up and supplemental 2013 Update weights; and

- the variability of the weights themselves and the degree to which the sums of the individual weights matched calibration totals.

Some of the specific quality control checks employed for nonresponse and calibration weight adjustments are described below.

Nonresponse weight adjustment quality control checks included the following:

- Weight sums after nonresponse adjustment matched weight sums before nonresponse adjustment. This assessment included the overall weight sum and weights sums by the levels of the categorical variables used in the nonresponse model, such as student race/ethnicity, sex, and base-year school type.

- Overall unequal weighting effect (UWE) after nonresponse adjustment was not substantially higher than the overall UWE prior to nonresponse adjustment. As a general rule of thumb, increases in the overall UWE were kept within 10 percent of the overall UWE prior to nonresponse adjustment.

- UWEs before nonresponse adjustment were computed for the main effect variables[34] used in the nonresponse models and compared to the corresponding UWEs after nonresponse adjustment. As a general rule of thumb, increases of 10 percent in the UWEs were considered acceptable.

Calibration weight adjustments quality control checks included the following:

- Weight sums after calibration were compared with target control totals to verify equivalence. The target control totals included totals defined for the school type, region, state (if part of the 10 state-representative public-school samples), and metropolitan status.

- UWEs were compared before calibration and compared with corresponding UWEs after calibration by school type, region, augmented state, metropolitan status, sex, and race.

[34] Main effect variables are variables such as sex and race/ethnicity which are distinguished from interacted variables.

Additional QC checks for BRR weight construction included the following:

- Comparing overall UWEs, minimum weights, maximum weights, and average weights across each set of 200 replicates to verify comparability of the replicate weight distributions.

6.4 Choosing an Analytic Weight

The choice about which weights to create for HSLS:09 data is driven by the need to maximize the analytic utility for the research community. Analyses may incorporate data obtained from a particular instrument within a round of the study (e.g., student questionnaire responses in the first follow-up) or combinations of data from multiple instruments across multiple rounds, such as student and parent questionnaire responses in the base year and first follow-up. As discussed in the *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015) and repeated here, weights were derived that incorporate many, but not all, possible combinations of data sources and rounds of data collection.[35]

The second follow-up data file contains a total of 25 analytic weights: five weights for analysis of the base-year data, four weights to be used in conjunction with the first follow-up data, six weights to be used for analysis involving the 2013 Update, five weights for analyses using High School Transcript data, and five weights for analyses of second follow-up data.

The analysis weights presented in table 29 can be used for analysis of data collected in a single study round or data collected across multiple study rounds. The weights designed to be used in analysis of a single round of data are classified as "single-round" weights, and the weights that may be used to analyze data collected from multiple study rounds are classified as "multiround."

Analyses of base-year data involving only the student assessment data or student questionnaire responses should use W1STUDENT, and base-year analyses that include parent responses from the base year should utilize W1PARENT. Analysis of school administrator or counselor responses, in the context of the HSLS:09 base-year school population, should utilize W1SCHOOL. Similarly, analyses involving only the first follow-up student questionnaire or assessment data should utilize W2STUDENT, and analyses involving first follow-up parent responses should utilize W2PARENT. Analyses involving only 2013 Update data should use

[35] The creation of additional HSLS:09 weights was considered. However, 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 data sources.

W3STUDENT for analyzing questionnaire responses, and analyses of only High School Transcript data should use W3HSTRANS. Analyses that involve only second follow-up questionnaire responses should use W4STUDENT.

Some of the analysis weights presented in table 29 are appropriate to use when analyzing data collected across multiple study rounds or from multiple data sources. For example, an analysis seeking to determine base-year predictors of on-time high school graduation should incorporate the analysis weight W3W1STUTR. Similarly, an analysis seeking to determine prior-round predictors of income as of the second follow-up should use W4W1W2W3STU if the set of possible predictors was limited to student questionnaire responses.

Table 29. HSLS:09 analytic weights

HSLS:09 round(s)	Universe[1]	Estimation	Variable name	Nonresponse-adjusted component(s) in each weight[2]
Base year	All study-eligible schools	Single-round	W1SCHOOL	BY School
Base year	All study-eligible students in base year[3]	Single-round	W1STUDENT	BY Student
			W1PARENT	BY Student * BY Parent
			W1SCITCH	BY Student * BY Science teacher
			W1MATHTCH	BY Student * BY Math teacher
First follow-up	9th-grade cohort[3]	Single-round	W2STUDENT	F1 Student
			W2PARENT	F1 Parent
Base year and first follow-up	9th-grade cohort[3]	Multiround	W2W1STU	BY/F1 Student
			W2W1PAR	BY/F1 Student * BY/F1 Parent
2013 Update	9th-grade cohort[3]	Single-round	W3STUDENT	U13 Student
Base year and 2013 Update	9th-grade cohort[3,4]	Multiround	W3W1STU	BY/U13 Student
			W3W1MATHTCH	BY/U13 Student * BY Math teacher[5]
			W3W1SCITCH	BY/U13 Student * BY Science teacher[6]
First follow-up and 2013 Update	9th-grade cohort[3,4]	Multiround	W3W2STU	F1/U13 Student
Base year, first follow-up, and 2013 Update	9th-grade cohort[3,4]	Multiround	W3W1W2STU	BY/F1/U13 Student
High School Transcript	9th-grade cohort[3,4]	Single-round	W3HSTRANS	High School Transcript
High School Transcript and 2013 Update	9th-grade cohort[3,4]	Multiround	W3STUDENTTR	High School Transcript * U13 Student
High School Transcript, base year, and 2013 Update	9th-grade cohort[3,4]	Multiround	W3W1STUTR	High School Transcript * BY/U13 Student

See notes at end of table.

Table 29. HSLS:09 analytic weights—Continued

HSLS:09 round(s)	Universe[1]	Estimation	Variable name	Nonresponse-adjusted component(s) in each weight[2]
High School Transcript, base year, first follow-up, and 2013 Update	9th-grade cohort[3,4]	Multiround	W3W1W2STUTR	High School Transcript * BY/F1/U13 Student
High School Transcript, first follow-up, and 2013 Update	9th-grade cohort[3,4]	Multiround	W3W2STUTR	High School Transcript * F1/U13 Student
Second follow-up	9th-grade cohort[3,4]	Single-round	W4STUDENT	F2 Student
Base year and second follow-up	9th-grade cohort[3,4]	Multiround	W4W1STU W4W1STUP1	BY/F2 Student BY/F2 Student * BY Parent[7]
Base year, first follow-up, and second follow-up	9th-grade cohort[3,4]	Multiround	W4W1STUP1P2	BY/F2 Student * BY/F1 Parent[8]
Base year, first follow-up, 2013 Update, and second follow-up	9th-grade cohort[3,4]	Multiround	W4W1W2W3STU	BY/F1/U13/F2 Student

[1] The sum of the associated analytic weights estimates the population count for the universe.
[2] Student-level weights are derived from the school analytic weight and therefore are also adjusted for school nonresponse. Unless otherwise specified, the weights were additionally adjusted for nonresponse within the specified round(s) of data collection.
[3] The subpopulation associated with the public-use file for student weights is restricted to 9th-grade students who were capable of participating in the student questionnaire and math assessment in the base year and first follow-up.
[4] Excludes those from the cohort who were deceased at the time of the latest data collection accounted for by the weight.
[5] Accounts for student nonresponse in the base year, nonresponse in the 2013 Update, and base-year math teacher nonresponse.
[6] Accounts for student nonresponse in the base year, nonresponse in the 2013 Update, and base-year science teacher nonresponse.
[7] Accounts for student nonresponse in the base year, nonresponse in the second follow-up, and base-year parent nonresponse.
[8] Accounts for student nonresponse in the base year, nonresponse in the second follow-up, base-year parent nonresponse, and first follow-up parent nonresponse.
NOTE: The symbol "*" should be interpreted as "and." For example, the W1PARENT weight was developed using adjustments for student and parent nonresponse. BY = base year, F1 = first follow-up, U13 = 2013 Update, and F2 = second follow-up.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

The number and percentage of completed surveys, High School Transcript responses, or their combinations for the student sample, and associated recommended weights for the HSLS:09 base-year, first follow-up, 2013 Update, High School Transcript, and second follow-up study rounds are summarized in table 30. Please note that, although the restricted-use file contains nonzero weights for students who were excluded from the base-year or first follow-up student survey because it was not offered in a format that allowed their meaningful participation (students referred to as "questionnaire incapable" in response status variables) in the base year or first follow-up, the weights for such students are set to zero in the corresponding public-use files. Inferences made using the public-use file therefore reflect a slightly smaller and slightly different set of students than that contained on the restricted-use file. To produce general population estimates that align with public-use estimates, set the restricted-use analysis weights to zero for any sample member classified as "questionnaire incapable" in the relevant base-year or first follow-up study. Two restricted-use variables, X1SQSTAT and X2SQSTAT, can be used to identify sample members classified as "questionnaire incapable" in the base year or first follow-up. If X1SQSTAT = 7, then a sample member was classified as "questionnaire incapable" in the base year, and if X2SQSTAT = 7, then a sample member was classified as "questionnaire incapable" in the first follow-up.

Table 30. Number and percentage of completed surveys, High School Transcript responses, or their combinations for the student sample, and associated recommended weights: Second follow-up

Study round and high school transcript combinations	Data source(s) and recommended weights	Eligible	Participated	Weighted percent[1]	Unweighted percent
Base year	BY Student questionnaire (W1STUDENT[2])	25,206	21,444	85.7	85.1
	BY Student assessment (W1STUDENT[3])	25,206	20,781	83.0	82.4
	BY Student and Parent questionnaires (W1PARENT[2])	25,206	16,429	65.3	65.2
	BY School administrator (W1STUDENT[3])	25,206	20,301	81.1	80.5
	BY School counselor (W1STUDENT[3])	25,206	19,505	77.7	77.4
	BY Teacher questionnaire[4]				
	Math teacher (W1MATHTCH[2])	23,621	16,035	65.1	67.9
	Science teacher (W1SCITCH[2])	22,597	14,629	63.6	64.7
First follow-up	F1 Student questionnaire (W2STUDENT[2])	25,184	20,594	82.0	81.8
	F1 Student assessment (W2STUDENT[3])	25,184	18,507	73.0	73.5
	F1 Parent questionnaire[5] (W2PARENT[2])	11,952	8,651	72.5	72.4
Base year and first follow-up	BY/F1 Student questionnaires (W2W1STU[2])	25,184	18,623	74.3	74.0
	BY/F1 Student assessments (W2W1STU[3])	25,184	16,356	64.7	65.0
	BY/F1 Student and Parent questionnaires[6] (W2W1PAR[2])	11,952	6,371	52.9	53.3
2013 Update	U13 Student questionnaire (W3STUDENT[2])	25,168	18,558	73.1	73.7
Base year and 2013 Update	BY/U13 Student questionnaires (W3W1STU[2])	25,168	17,117	67.6	68.0
	BY/U13 Student and BY Teacher questionnaires				
	Math teacher[7] (W3W1MATHTCH[2])	23,587	12,812	51.4	54.3
	Science teacher[8] (W3W1SCITCH[2])	22,566	11,803	50.7	52.3
First follow-up and 2013 Update	F1/U13 Student questionnaires (W3W2STU[2])	25,168	17,282	68.0	68.7
Base year, first follow-up, and 2013 Update	BY/F1/U13 Student questionnaires (W3W1W2STU[2])	25,168	15,857	62.5	63.0

See notes at end of table.

Table 30. Number and percentage of completed surveys, High School Transcript responses, or their combinations for the student sample, and associated recommended weights: Second follow-up—Continued

Study round and high school transcript combinations	Data source(s) and recommended weights	Eligible	Participated	Weighted percent[1]	Unweighted percent
High School Transcript	High School Transcript (W3HSTRANS[2])	25,167	21,928	87.7	87.1
High School Transcript and 2013 Update	High School Transcript and U13 Student questionnaire (W3STUDENTTR[2])	25,167	17,656	70.2	69.63
High School Transcript, base year, and 2013 Update	High School Transcript and BY/U13 Student questionnaires (W3W1STUTR[2])	25,167	16,303	64.7	64.4
High School Transcript, first follow-up, and 2013 Update	High School Transcript and F1/U13 Student questionnaires (W3W2STUTR[2])	25,167	16,525	65.6	64.9
High School Transcript, base year, first follow-up, and 2013 Update	High School Transcript and BY/F1/U13 Student questionnaires (W3W1W2STUTR[2])	25,167	15,188	60.4	59.8
Second follow-up	F2 Student questionnaire (W4STUDENT[2])	25,123	17,335	67.9	69.0
Second follow-up and base year	BY/F2 Student questionnaires (W4W1STU[2])	25,123	15,909	62.5	63.3
	BY/F2 Student and BY Parent questionnaires[9] (W4W1STUP1[2])	25,123	12,888	50.1	51.3
Second follow-up, base year, and first follow-up	BY/F2 Student and BY/F1 Parent questionnaires[5,10] (W4W1STUP1P2[2])	11,927	5,427	44.6	45.5
Second follow-up, base year, first follow-up, and 2013 Update	BY/F1/U13/F2 Student questionnaires (W4W1W2W3STU[2])	25,123	13,283	52.0	52.9

[1] All weighted percentages are calculated using the student base weight.
[2] Recommended weight, constructed to account for response to the data source.
[3] Recommended weight, not constructed specifically for response to the data source.
[4] Results for the math teacher questionnaire reflect students who were enrolled in a mathematics course in the base year; results for the science teacher questionnaire reflect students who were enrolled in a science course in the base year.
[5] Details of the parent subsample design are provided in section 3.3.4 of the *HSLS:09 Base Year to First Follow-up Data File Documentation* (Ingels et al. 2013).
[6] 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.
[7] Only sampled students who participated in both the base year and 2013 Update with a responding base-year math teacher are considered participants.
[8] Only sampled students who participated in both the base year and 2013 Update with a responding base-year science teacher are considered participants.
[9] Only sampled students who participated in both the base year and second follow-up with a responding parent in the base year are considered participants.
[10] Only sampled students who participated in both the base year and second follow-up with a responding parent in the base year and first follow-up are considered participants.
NOTE: All counts and computed rates are at the student level. BY = base year, F1 = first follow-up, U13 = 2013 Update, and F2 = second follow-up.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Choosing a weight for analyses can be complicated. To help in choosing a weight, researchers should first think in terms of the particular time period or data source of interest for the HSLS:09 population of students—base year, first follow-up, 2013 Update, High School Transcript, second follow-up, or some combination thereof. Next, researchers should consider the magnitude of nonresponse with the records included in the analyses and the associated nonresponse adjustment(s) for each weight.

As an example of how nonresponse magnitude might influence an analyst's decisions regarding which weight to use, consider a regression-based analysis. Records are excluded from a regression model if model covariates are missing, if the analysis weight is zero, or both. Consider an example in which 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 in the respective contextual data sources (i.e., parent and science teacher nonresponse, respectively). However, because neither addresses 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., containing nonmissing 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.

In the event that no weight accommodates interview data from all time periods and data sources of interest, researchers will have to assess the available weights to determine which weight should be used. A general rule of thumb is to select the weight that accounts for as many components of nonresponse as possible and, in the event of a tie, to select the weight that yields the most records for the analysis of interest. For illustration, suppose an analysis will use High School Transcript data and interview data from the second follow-up and base-year time periods. There is no analysis weight that explicitly accounts for interview nonresponse in the base year, for interview nonresponse in the second follow-up, and for missing High School Transcript data. However, there is one weight that accounts for two of the three sources of nonresponse, W4W1STU—this weight is therefore recommended for the analysis.

The HSLS:09 base year and first follow-up included two sources of contextual information that were not obtained in the 2013 Update or second follow-up:

9th-grade science and mathematics teacher interviews in the base year and parent data in the base year and first follow-up. In the base year, interviews with the science teacher and mathematics teacher were conducted for students taking the associated course in the 9th grade. Prior to the second follow-up, there were no weights created that specifically integrated a round after the base year with the teacher data collected in the base year. The supplemental 2013 Update weights were designed specifically to fill this void. There are now two weights that account for student nonresponse in the base year and 2013 Update, in conjunction with math teacher nonresponse: W3W1MATHTCH, which accounts for 9th-grade mathematics teacher nonresponse, and W3W1SCITCH, which accounts for 9th-grade science teacher nonresponse.

A note on incorporation of base-year teacher interview data into analyses. 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 9th grade and the end of 11th grade, but teacher characteristics were only measured for the fall of 9th 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.

As was stressed in the *HSLS:09 Base Year Data File Documentation* (Ingels et al. 2011), the teacher sample does *not* constitute a nationally representative or school-representative sample of 9th-grade mathematics and science teachers. The two separate mathematics and science teacher samples 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 datasets 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 a teacher-level 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.

If base-year teacher data are used in conjunction with data from other time periods or from noninterview sources, the premise in selecting a weight as discussed above applies. Consider an example in which 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 associated with first follow-up student data but not for the nonresponse 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, 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, compared to using the data and weights in their published state.

6.4.1 Base-year School-level Analysis

School-level analysis is only appropriate with the base-year school-level data. The HSLS:09 study design supports national estimates of schools with both 9th- and 11th-grade students in the base year of the survey, the 2009–10 academic year.[36]

- 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 to analyze school-level data, school administrator survey data, and counselor survey data, individually or in combination. Note that weighted values generated from school administrator and counselor responses provide information for the HSLS:09 target population of schools, not administrators or counselors.[37] Additional information on the construction of the school weight is provided in the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011).

6.4.2 Base-year Student-level Analysis

Analyses that include only base-year student-level data—no 2013 Update, High School Transcript, or second follow-up data—should utilize one of the four weights discussed in this section. To analyze base-year data combined with data from another round, use one of the weights discussed in sections 6.4.3, 6.4.4, and 6.4.5.

---

[36] 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 *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011).

[37] Questionnaire responses were requested from the lead counselor or counselor most knowledgeable about 9th-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.

- W1STUDENT. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in the base year. All records for sample students who participated in the base year have a positive (nonzero) weight. Estimates generated with this base-year student weight represent the HSLS:09 target population of 9th-grade students. This weight can be used to analyze base-year student assessment scores or survey data alone or in combination with the school, administrator, or counselor data.[38]

- W1PARENT. This weight accounts for school nonresponse, student questionnaire nonresponse, and parent nonresponse in the base year. All records for sample students who participated in the base year by completing the student questionnaire and whose parent who also participated in the base year have a positive (nonzero) weight. Estimates generated with this base-year student home-life weight are associated with the HSLS:09 target population of 9th-grade students. This weight can be used for the analysis of base-year parent data alone or in conjunction with base-year student, school, administrator, or counselor data.

- W1SCITCH. This weight includes nonresponse adjustments for (1) school nonresponse, (2) student questionnaire 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 have a positive (nonzero) weight. Estimates generated with this base-year science-course enrollee weight are associated with the subgroup of 9th-grade students in the HSLS:09 target population taking a science course in the 9th grade. These estimates *do not* reflect the population of all science teachers of 9th-grade students because science teachers themselves were not sampled directly; see sections 3.4 and 6.5 of the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011) for further information. This weight can be used for the analysis of science teacher data in conjunction with base-year student, school, administrator, or counselor data.

- W1MATHTCH. This weight includes nonresponse adjustments for (1) school nonresponse, (2) student questionnaire nonresponse, and

[38] An analysis of the nonresponse in the combined student and administrator or counselor data did not indicate the need for additional student-level weights.
[39] The sum of the weights estimates the total number of 9th-grade students in the HSLS:09 target population taking a science course and is less than the total number of 9th-grade students.

(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 have a positive (nonzero) weight. Estimates generated with this base-year mathematics-course enrollee weight are associated with the subgroup of 9th-grade students in the HSLS:09 target population taking a mathematics course in the 9th grade. These estimates *do not* reflect the population of all mathematics teachers of 9th-grade students because mathematics teachers themselves were not sampled directly; see sections 3.4 and 6.5 of the *HSLS:09 Base-Year Data File Documentation* (Ingels et al. 2011) for further information. This weight can be used for the analysis of mathematics teacher data in conjunction with base-year student, school, administrator, or counselor data.

6.4.3 First Follow-up Student-level Analysis

Four weights were constructed to be used in analysis of first follow-up data. Two weights, W2STUDENT and W2PARENT, were constructed to be used in analysis of first follow-up data and two other weights, W2W1STU and W2W1PAR, were constructed to be used in analysis of only first follow-up data in conjunction with base-year data. If a researcher is analyzing data only from the first follow-up, one of the first two first follow-up weights discussed in this section should be used (W2STUDENT or W2PARENT). If data from the first follow-up and the base year are to be analyzed then one of the two other weights, W2W1STU and W2W1PAR, should be used.

- W2STUDENT. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire 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 have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of 9th-grade students.[41] 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 data,

[40] The sum of the weights estimates the total number of 9th-grade students in the HSLS:09 target population taking a mathematics course and is less than the total number of 9th-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 homeschooled students and nonresponding transfer students.

administrator/counselor data from either round of HSLS:09, or teacher data from the base year.[42,43]

- 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.[44,45] All records for sample students with a parent who participated in the first follow-up have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of 9th-grade students. This weight can be used for analysis of first follow-up parent responses alone or in combination with student questionnaire data, assessment data, or both; administrator/counselor data from either round of HSLS:09; or teacher data from the base year.[46]

- W2W1STU. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire 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 have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students. This weight can be used for analyses incorporating both base-year and first follow-up student questionnaire or assessment data; alone or in combination with: administrator/counselor data from the base year, the first follow-up, or both; or teacher data from the base year.[47]

---

[42] Not all students were taking science or mathematics courses in the 9th grade. Therefore, analyses involving the base-year teacher responses will provide estimates for the subgroup of 9th-grade students in the HSLS:09 target population taking the associated course.
[43] Note that estimates generated with first follow-up student data and W2STUDENT in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[44] Note that W2PARENT differs slightly from the base-year weight W1PARENT. Unlike in 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.
[45] Note that student data are not available for 355 student records that have first follow-up parent data because of student nonresponse in the first follow-up.
[46] Note that estimates generated with first follow-up student data and W2PARENT in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[47] Note that estimates generated with first follow-up student data and W2W1STU in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

- W2W1PAR. This weight accounts for (1) school nonresponse in the base year, (2) student questionnaire nonresponse in the base year and the first follow-up, (3) subsampling of parents for the first follow-up, and (4) parent nonresponse in 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 have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students. This weight can be used for analysis incorporating base-year and first follow-up home-life (contextual) data obtained from the parent questionnaires, alone or in combination with student questionnaire data, assessment data, or both; administrator/counselor data from either round of HSLS:09; or teacher data from the base year.[48]

6.4.4 2013 Update and High School Transcript Student-level Analysis

Eleven weights were constructed to account for nonresponse to the 2013 Update questionnaire or nonresponse associated with High School Transcript data collection. The nine 2013 Update and High School Transcript weights developed as part of the 2013 Update and High School Transcript rounds are described in section 6.4.4.1. Two supplemental 2013 Update weights, W3W1MATHTCH and W3W1SCITCH, were constructed as part of the second follow-up to better support analysis of 2013 Update data in conjunction with base-year teacher data. These two supplemental weights are described in section 6.4.4.2.

6.4.4.1 2013 Update and High School Transcript weights

The nine weights constructed as part of the 2013 Update and High School Transcript data collection are designed for use in analysis of 2013 Update questionnaire data and data from the High School Transcript collection, either alone or in conjunction with each other or with data from other study rounds.

Analyses that involve only data collected from the 2013 Update questionnaire should use W3STUDENT, and analyses that involve only data collected in the High School

[48] Note that estimates generated with first follow-up student data and W2W1PAR in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

Transcript study should use W3HSTRANS. Descriptions of these two weights and the other 2013 Update and High School Transcript weights are provided below.

- W3STUDENT. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in the 2013 Update only, regardless of the student's response status in other rounds. All records for sample students who participated in the 2013 Update have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of 9th-grade students who are alive as of the 2013 Update.[49] This weight can be used for the analysis of 2013 Update survey data, alone or in combination with the school characteristics data or administrator/counselor data from HSLS:09.[50]

- W3W1STU. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in both the base year and the 2013 Update. All records for sample students who participated in the base year and 2013 Update have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the 2013 Update. This weight can be used for analyses that examine student data from the base year and the 2013 Update, alone or in combination with administrator/counselor data.[51]

- W3W2STU. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in the first follow-up and the 2013 Update. All records for sample students who participated in the first follow-up and 2013 Update have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the 2013 Update. This weight can be used for analysis of student data from the first follow-up and

[49] Weights were computed for deceased students and then subsequently removed such that the sum of the weights is representative of the HSLS:09 target population of 9th-grade students who were alive as of the 2013 Update data collection. This method was implemented for all six 2013 Update student weights.

[50] Note that estimates generated with 2013 Update student data and W3STUDENT in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

[51] Note that estimates generated with 2013 Update student data and W3W1STU in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

the 2013 Update, alone or in combination with administrator/counselor data from the first follow-up of HSLS:09.[52]

- W3W1W2STU. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in the base year, first follow-up, and 2013 Update. All records for sample students who participated in the base year, first follow-up, and 2013 Update have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the 2013 Update. This weight can be used for analysis of student data from the base year, the first follow-up, and the 2013 Update, alone or in combination with administrator/counselor data from the base year or the first follow-up.[53]

- W3HSTRANS. This weight accounts for (1) base-year school nonresponse and (2) High School Transcript nonresponse only, regardless of the student's response status in other data collection rounds of HSLS:09. All records for sample students for whom a high school transcript was collected have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of 9th-grade students who are alive at the time of High School Transcript data collection.[54] This weight can be used for the analysis of High School Transcript data, alone or in combination with school characteristics data, administrator/counselor data, or teacher data from the base year.[55]

- W3STUDENTTR. This weight accounts for (1) base-year school nonresponse, (2) High School Transcript nonresponse, and (3) student questionnaire nonresponse in the 2013 Update only, regardless of the student's response status in other rounds. All records for sample students for whom a high school transcript was collected *and* who participated in the

[52] Note that estimates generated with 2013 Update student data and W3W2STU in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[53] Note that estimates generated with 2013 Update student data and W3W1W2STU in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[54] Weights were computed for deceased students and then subsequently removed such that the sum of the weights is representative of the HSLS:09 target population of 9th-grade students who were alive as of the High School Transcript data collection. This method was implemented for all five High School transcript weights.
[55] Note that estimates generated with High School Transcript data and W3HSTRANS in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

2013 Update have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of 9th-grade students who are alive at the time of High School Transcript data collection. This weight can be used for the analysis of High School Transcript data in conjunction with 2013 Update student data, alone or in combination with school characteristics data, administrator/counselor data, or teacher data from the base year.[56]

- W3W1STUTR. This weight accounts for (1) base-year school nonresponse, (2) High School Transcript nonresponse, and (3) student questionnaire nonresponse in both the base year and the 2013 Update. All records for sample students for whom a high school transcript was collected *and* who participated in the base year *and* 2013 Update have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive at the time of High School transcript data collection. This weight can be used for analyses that examine student data from the base year and the 2013 Update and incorporate the High School Transcript data, alone or in combination with administrator/counselor data, teacher data from the base year, or both.[57]

- W3W2STUTR. This weight accounts for (1) base-year school nonresponse, (2) High School Transcript nonresponse, and (3) student questionnaire nonresponse in both the first follow-up and the 2013 Update. All records for sample students for whom a high school transcript was collected and who participated in the first follow-up and 2013 Update have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive at the time of High School Transcript data collection. This weight can be used for analysis of student data from the first follow-up and the 2013 Update that incorporate High School Transcript data, alone or in combination with administrator/counselor data from the first follow-up of HSLS:09.[58]

---

[56] Note that estimates generated with 2013 Update and High School Transcript data and W3STUDENTTR in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

[57] Note that estimates generated with High School Transcript data and W3W1STUTR in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

[58] Note that estimates generated with High School Transcript data and W3W2STUTR in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

- W3W1W2STUTR. This weight accounts for (1) base-year school nonresponse, (2) High School Transcript nonresponse, and (3) student questionnaire nonresponse in the base year, first follow-up, and 2013 Update. All records for sample students for whom a high school transcript was collected *and* who participated in the base year, first follow-up, *and* 2013 Update have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive at the time of High School Transcript data collection. This weight can be used for analysis of student data from the base year, the first follow-up, and the 2013 Update that incorporate High School Transcript data, alone or in combination with: administrator/counselor data from the base year, the first follow-up of HSLS:09, or both; or teacher data from the base year.[59]

6.4.4.2 Supplemental 2013 Update weights constructed in the second follow-up

- W3W1MATHTCH. This weight accounts for (1) base-year school nonresponse, (2) student questionnaire nonresponse in both the base year and the 2013 Update, and (3) math teacher nonresponse in the base year. All records for sample students who participated in the base year and 2013 Update with a responding base-year math teacher have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students enrolled in a base-year math course who are alive as of the 2013 Update. This weight can be used for analysis of student data from the base year and the 2013 Update with base-year math teacher data, alone or in combination with administrator/counselor data.

- W3W1SCITCH. This weight accounts for (1) base-year school nonresponse, (2) student questionnaire nonresponse in both the base year and the 2013 Update, and (3) science teacher nonresponse in the base year. All records for sample students who participated in the base year and 2013 Update with a responding base-year science teacher have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students enrolled in a base-year science course who are alive as of the 2013 Update. This weight can be used for analysis of student data from the base year and the 2013 Update with

[59] Note that estimates generated with High School Transcript data and W3W1W2STUTR in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

base-year science teacher data, alone or in combination with administrator/counselor data.

6.4.5 Second Follow-up Student-level Analysis

There are five student weights constructed to be used in analysis of second follow-up student questionnaire responses. One weight, W4STUDENT, is designed to be used in analysis of only second follow-up data, and two weights, W4W1STU and W4W1W2W3STU, are designed for analysis of second follow-up questionnaire data in conjunction with student questionnaire data from other study rounds. Two other weights, W4W1STUP1 and W4W1STUP1P2, were constructed to be used in analysis of second follow-up data in conjunction with base-year student questionnaire and parent questionnaire data and first follow-up parent questionnaire data.

- W4STUDENT. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in the second follow-up only, regardless of student response status in other rounds. All records for sample students who participated in the second follow-up have a positive (nonzero) weight. The estimates generated with these weights are associated with the HSLS:09 target population of 9th-grade students who are alive as of the second follow-up.[60] This weight can be used for the analysis of second follow-up survey data, alone or in combination with school characteristics data or administrator/counselor data from any round of HSLS:09, teacher data from the base year, or both.[61]

- W4W1STU. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in both the base year and the second follow-up. All records for sample students who participated in the base year and second follow-up have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the second follow-up. This weight can be used for analysis of student data from the base year and

[60] Weights were computed for deceased students and then subsequently removed such that the sum of the weights is representative of the HSLS:09 target population of 9th-grade students who are alive as of the second follow-up data collection. This method was implemented for all five second follow-up student weights.

[61] Note that estimates generated with second follow-up student data and W4STUDENT in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

the second follow-up, alone or in combination with administrator/counselor data.[62]

- W4W1W2W3STU. This weight accounts for (1) base-year school nonresponse and (2) student questionnaire nonresponse in the base year, first follow-up, 2013 Update, and second follow-up. All records for sample students who participated in the base year, first follow-up, 2013 Update, and second follow-up have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the second follow-up. This weight can be used for analysis of student data from the base year, the first follow-up, the 2013 Update, and the second follow-up, alone or in combination with administrator/counselor data from the base year and/or the first follow-up.[63]

- W4W1STUP1. This weight accounts for (1) base-year school nonresponse, (2) student questionnaire nonresponse in both the base year and the second follow-up, and (3) parent nonresponse in the base year. All records for sample students who participated in the base year and second follow-up with a responding parent in the base year have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the second follow-up. This weight can be used for analysis of student data from the base year and the second follow-up with base-year parent data, alone or in combination with administrator/counselor data.[64]

- W4W1STUP1P2. This weight accounts for (1) base-year school nonresponse, (2) the parent subsampling in the first follow-up,[65] (3) student questionnaire nonresponse in both the base year and the second follow-up, and (4) parent nonresponse in both the base year and first follow-up. All records for sample students who participated in the base year and second

[62] Note that estimates generated with second follow-up student data and W4W1STU in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[63] Note that estimates generated with second follow-up student data and W4W1W2W3STU in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[64] Note that estimates generated with second follow-up student data and W4W1STUP1 in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.
[65] Details of the parent subsample design are provided in section 3.3.4 of the *HSLS:09 Base Year to First Follow-up Data File Documentation* (Ingels et al. 2013).

follow-up with a responding parent in the base year and first follow-up have a positive (nonzero) weight. The estimates generated with this weight are associated with the HSLS:09 target population of 9th-grade students who are alive as of the second follow-up. This weight can be used for analysis of student data from the base year and the second follow-up, with base-year parent data and first follow-up parent data, alone or in combination with administrator/counselor data.[66] Researchers interested in evaluating base-year and second follow-up student data with first follow-up parent data, but without base-year parent data, are advised to use this weight for analyses because the results will still be representative of the HSLS:09 target population.

6.5 Measures of Precision: Standard Errors and Design Effects

This section discusses the standard errors and design effects associated with HSLS:09 estimation. Readers may refer to appendix I for tables providing survey estimates, standard errors, and design effects for various domains of interest, computed using the primary second follow-up weight.

6.5.1 Standard Errors

Complex sample designs, like that used for HSLS:09, result in data that violate the assumptions that are normally required to assess the statistical significance of results. The standard errors of the estimates from complex surveys may vary from those that would be expected if the sample were a simple random sample and the observations were independent. Some standard software packages, however, do not calculate standard error estimates that account for complex sampling design used to select the school and student samples. This incorrect design assumption can lead to estimated variances and confidence intervals that are too small, which may lead to incorrect results from hypothesis tests. Variables have been created for HSLS:09 to support two methods of standard error estimation that account for the HSLS:09 complex sample design: (1) a balanced repeated replication (BRR) variance estimation method using the BRR weights and the associated analytic weight and (2) a linearization variance estimation method through a Taylor series approximation using analytic weights and variables that represent school sampling strata and primary sampling

[66] Note that estimates generated with second follow-up student data and W4W1STUP1P2 in conjunction with the base-year teacher responses are no longer representative of the HSLS:09 target population of 9th-grade students and should be used with caution.

units.[67] Please note that variables to support these two methods of variance estimation are available to users of the restricted-use data, but only the BRR variance estimation method is supported for users of public-use data. Researchers are advised to use specialized software such as SUDAAN, SAS, or Stata that adjusts standard errors to account for the complex sampling design using one of these methods. Examples of code for these software programs are provided below.

The importance of correct variance estimation is further emphasized in this section through a discussion of the BRR and linearization methodologies.

The two methods of variance estimation supported through available HSLS:09 variables are BRR and Taylor series linearization. BRR variance estimation is supported with either the HSLS:09 restricted-use or public-use files. This method does not need the analytic stratum and primary sampling unit (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 Data File Documentation* (Ingels et al. 2011), *Base Year to First Follow-Up Data File Documentation* (Ingels et al. 2013), and *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015), the replicate weights account for unequal selection probabilities, stratification, and clustering; incorporate nonresponse and calibration adjustments; and produce standard error estimates that are in general slightly larger than the corresponding estimates calculated with linearization (Wolter 2007).

To create the BRR weights, the original sampling strata were collapsed into 199 BRR strata with representation across the characteristics used in school 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 as follows:

[67] NCES statistical standards recommend the use of replicate variance estimation over linearization methods. The sample design variables, strata, and primary sampling units were suppressed from the public-use file as one measure of disclosure avoidance (see section 6.9 for information regarding the disclosure risk analysis and protection).

$$var\left(\hat{\theta}\right) = \frac{1}{200}\sum_{a=1}^{200}\left(\hat{\theta}_{(a)} - \hat{\theta}\right)^2 \quad (6\text{-}1)$$

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 ($a = 1, \ldots, 200$).

Taylor series linearization variance estimation requires software that uses the analytic weight, analytic stratum, and PSU identifiers to compute standard errors that are adjusted to account for the complex sample design (see, e.g., Binder [1983]; Woodruff [1971]). The PSU and stratum identifiers are provided in two restricted-use variables, PSU and STRAT_ID. The PSU variable contains a unique value randomly generated for each sampled school. The 450 values of STRAT_ID were constructed in the base year by combining two to three schools into one analysis stratum in such a way as to maximize retention of the original two-stage sample design and also increase the precision of the estimates through the degrees of freedom (Chromy 1981). To lower disclosure risk, variables to support linearization variance estimation are only provided through the HSLS:09 restricted-use file, which, unlike the public-use file, contains the stratum and PSU variables.

Currently available software that can compute standard errors adjusted to account for a complex sample design includes SUDAAN,[68] SAS SURVEY procedures,[69] WesVar,[70] Stata,[71] R,[72] and SPSS.[73] Example SAS-callable SUDAAN code for producing estimated means and standard errors using the linearization and BRR methods are shown in figures 10 and 11, respectively. The corresponding Stata code is provided in figures 12 and 13, and SAS code provided in figures 14 and 15.

[68] See http://sudaansupport.rti.org/sudaan/index.cfm.
[69] See the most recent *SAS User's Guide*, located at http://support.sas.com/documentation/.
[70] See https://www.westat.com/our-work/information-systems/wesvar-support.
[71] See https://www.stata.com/.
[72] See https://www.r-project.org/.
[73] See https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software.

Figure 10. Example SAS-callable SUDAAN code to calculate an estimated mean and linearization standard error for a second follow-up 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 W4STUDENT;                        *Main analytic weight;
  VAR <analysis variable>;                    *Analysis variable;
  PRINT MEAN SEMEAN / STYLE=NCHS;     *Mean and standard error;
RUN;
```

Figure 11. Example SUDAAN code to calculate an estimated mean and replicate (BRR) standard error for a second follow-up student-level longitudinal analysis

```
PROC DESCRIPT DATA=<filename> DESIGN=BRR;
  WEIGHT W4STUDENT;                         *Main analytic weight;
  REPWGT W4STUDENT001-W4STUDENT200;     *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.

Figure 12. Example Stata code to calculate an estimated mean and linearization standard error for a second follow-up student-level analysis

```
SVYSET PSU [PWEIGHT=W4STUDENT], STRATA (STRAT_ID) VCE(LINEAR),
singleunit(centered)

SVY, SUBP (<domain variable >) : MEAN < analysis variable >
```

Figure 13. Example Stata code to calculate an estimated mean and replicate (BRR) standard error for second follow-up student-level analysis

```
SVYSET [PWEIGHT=W4STUDENT], BRRWEIGHT(W4STUDENT001-
W4STUDENT200) VCE(BRR) MSE

SVY, SUBP (<domain variable >) : MEAN < analysis variable >
```

NOTE: BRR = balanced repeated replication.

Figure 14. Example SAS code to calculate an estimated mean and linearization standard error for a second follow-up student-level analysis

```
PROC SURVEYMEANS DATA=<filename> VARMETHOD=TAYLOR NOMCAR;
  STRATA STRAT_ID;                              *Analysis stratum;
  CLUSTER PSU;                                      *Analysis PSU;
  DOMAIN (<domain variable>);            *Subset to reporting domain;
  WEIGHT W4STUDENT;                            *Main analytic weight;
  VAR <analysis variable>;                        *Analysis variable;
RUN;
```

Figure 15. Example SAS code to calculate an estimated mean and replicate (BRR) standard error for a second follow-up student-level longitudinal analysis

```
PROC SURVEYMEANS DATA=<filename> VARMETHOD=BRR;
  WEIGHT W4STUDENT;                           *Main analytic weight;
  REPWEIGHTS W4STUDENT001-W4STUDENT200;  *BRR replicate
  weights;
  DOMAIN (<domain variable>);             *Subset to reporting domain;
  VAR <analysis variable>;                          *Analysis variable;
RUN;
```

NOTE: BRR = balanced repeated replication.

Standard errors for a select number of variables are provided in appendix I along with their design effects, which are 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,

$$deff = \frac{\hat{V}_d\left(\hat{\theta}\right)}{\hat{V}_s\left(\hat{\theta}\right)}, \tag{6-2}$$

for an estimated 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 sample size.

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

$$deft = \sqrt{\frac{\hat{V}_d\left(\hat{\theta}\right)}{\hat{V}_s\left(\hat{\theta}\right)}}, \tag{6-3}$$

where the components are the same as defined for expression (6-2).

As noted in section 6.5.1, correct estimation of the variance of estimates requires the use of specialized software that can account for unequal selection probabilities, stratification, and clustering. In situations where software is unable to adjust for stratification and clustering but can accommodate weights, design effects may be used to approximate design-based variance and standard error estimates and thereby to produce associated test statistics that account for the estimated design-based variance.

The first step in approximating design-based variance estimates requires construction of normalized analysis weights. Given one of the analysis weights, *w*, defined in section 6.4, normalized analysis weights are defined as

$$w_{i,norm} = w_i * \frac{n}{\sum_{i=1}^{n} w_i} \tag{6-4}$$

where *n* corresponds to the number of observations with a positive weight, *i* indexes the set of respondents with a positive weight, and $\sum_{i=1}^{n} w_i$ is the sum of the analysis weights.

There are three methods that may be used to produce *t* and *F* test statistics using approximated design-based variance estimates. The first method involves approximating the design-based variance estimate and using it to manually calculate the test statistics. In this first method, the normalized weights are used to estimate the simple random sampling variance or standard error of the estimator of interest using the available software. The design-based variance estimate may be approximated by multiplying the variance estimate produced from the software by an appropriate value of *deff*. Symbolically,

$$\hat{V}_d(\hat{\theta}) \sim \hat{V}_S(\hat{\theta}) * deff \tag{6-5}$$

where $\hat{V}_s\left(\hat{\theta}\right)$ is provided by the software, and *deff* may correspond to a specific estimate or may be the median or mean of *deff* over several estimates. If the estimate of interest is for a subpopulation, then the value used for *deff* may be generated from a subgroup of respondents. The design effects reported in table 31 and those provided in appendix I may also be used for this second step. The approximate design-based variance estimates may be used to manually compute *t* and *F* test statistics.

The second method involves using the available software along with the normalized weights to generate *t* and *F* test statistics and then dividing the *t* statistic by an appropriate *deft* value and dividing the *F* statistics by an appropriate *deff* value.

The third method requires computing a new analysis weight by dividing the normalized weights by an appropriate value of *deff* and using this new analysis weight with the available software, using the test statistics produced with the software for inference.

The HSLS:09 second follow-up *deff*/*deft* analysis included 37 variables associated with the second follow-up survey. 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* estimates were calculated using a model-based formulation, corresponding to the deff4 option in SUDAAN. As in the first follow-up and 2013 Update, the items were chosen using two criteria: (1) variables common to the HSLS:09 prior rounds' design effect analysis; and (2) variables included in several other NCES studies such as the Education Longitudinal Study of 2002 (ELS:2002) and the National Education Longitudinal Study of 1988 (NELS:88). The *deff* and *deft* estimates are provided in appendix I for the 37 second follow-up survey items chosen using the above-specified criteria. The average *deff* and *deft* across the 37 items is presented in table 31.

Table 31. Average design effects (*deff*) and root design effects (*deft*) for second follow-up student variables

		Final student weight	
Characteristic[1]	Student respondents	Average *deff*[2]	Average *deft*[3]
Total	**17,335**	**4.1**	**2.0**
School type			
Public	14,103	3.7	1.9
Private	3,232	4.2	2.0
Region			
Northeast	2,724	5.8	2.3
Midwest	4,694	2.8	1.7
South	6,969	3.3	1.8
West	2,948	4.1	2.0
Locale			
City	5,034	5.9	2.4
Suburban	6,311	3.0	1.7
Town	1,967	3.2	1.7
Rural	4,023	3.1	1.7
Student sex			
Male	8,464	3.5	1.9
Female	8,871	3.7	1.9
Student race/ethnicity[4]			
Hispanic	2,712	4.3	2.0
Asian	1,477	5.2	2.2
Black	1,779	3.1	1.7
Other	11,367	2.8	1.7
Socioeconomic status[5]			
Low SES	2,623	3.6	1.9
Middle SES	9,977	3.3	1.8
High SES	4,702	2.8	1.6

[1] The school characteristics (school type, region, and locale) presented here reflect the information obtained during the HSLS:09 base year and do not contain updated information presented on the cumulative data file. The demographic characteristics (sex, race/ethnicity, and socioeconomic status) presented here reflect information obtained during the HSLS:09 base year and updated in the first follow-up; these demographics were not updated in the 2013 Update or second follow-up rounds of sampling.
[2] The formula for the design effect (*deff*) is provided in expression (6-2).
[3] The formula for the root design effect (*deft*) is provided in expression (6-3).
[4] Race/ethnicity as defined in the student questionnaire. Race categories exclude persons of Hispanic ethnicity.
[5] Categories for socioeconomic status (SES) were defined using the SES quintile variable from the first follow-up (X2SESQ5), where X2SESQ5 = 1 (1st quintile) represents low SES, X2SESQ5 = 5 (5th quintile) represents high SES, and the three middle quintiles were classified as middle SES.
NOTE: Design effects and standard errors computed using the W4STUDENT weight.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

6.6 Item-level Declined Response

This section presents an analysis of the rate at which respondents declined to answer questions that they were asked in the survey: the item-level declined response rate. The rate of item-level declined response is a data quality measure used to identify troublesome interview items and better understand the experiences of sample members in completing the interview. This declined response analysis excludes items not answered because (1) the respondent was routed around the item, or (2) the item was not included on the abbreviated survey to which the sample member responded, or (3) the respondent exited the interview and missed items after the breakoff point. Therefore, the rate of item-level declined response is defined as an item-level quotient in which the numerator is the number of -9 values, defined as *Item-missing, nonresponse*, for a given item and the denominator is the number of respondents administered the item. This definition therefore produces the fraction of items that are not answered within a survey when the respondent was eligible for the item, which differs from the item-nonresponse rates used for nonresponse bias analysis, as reported in section 6.7.

Items that were administered to at least 50 respondents and had declined response rates of 5 percent or greater are presented in table 32 and discussed in this section. Overall, the item-level declined response analysis identified 11 items that had more than 5 percent missing data.

Table 32. Item-level declined response over 5 percent, by variable and mode: 2016

	All modes			Telephone[1]			Web[2]			Field[3]		
Item	Declined (*n*)	Admin (*n*)	Percent	Declined (*n*)	Admin (*n*)	Percent	Declined (*n*)	Admin (*n*)	Percent	Declined (*n*)	Admin (*n*)	Percent
S4INCOMESPS	84	639	13.1	18	140	12.9	64	482	13.3	2	17	11.8
S4INCOME	1,963	15,457	12.7	565	2,652	21.3	1,369	12,589	10.9	29	216	13.4
S4MLTENDY	10	79	12.7	0	25	0.0	10	54	18.5	—	—	—
S4MLTENDM	8	79	10.1	0	25	0.0	8	54	14.8	—	—	—
S4UNEMP16FB	400	5,157	7.8	38	805	4.7	355	4,282	8.3	7	70	10.0
S4JOBDUTY2	1,131	15,004	7.5	177	2,788	6.3	937	11,947	7.8	17	269	6.3
S4OCC30EARN	1,126	15,304	7.4	461	2,659	17.3	657	12,436	5.3	8	209	3.8
S4EVERAPPLY	340	4,645	7.3	38	1,097	3.5	301	3,397	8.9	1	151	0.7
S4PRVLOANAMT	108	1,861	5.8	22	181	12.2	86	1,675	5.1	0	5	0.0
S4WORKENDM	206	3,771	5.5	39	634	6.2	166	3,086	5.4	1	51	2.0
S4WORKENDY	200	3,771	5.3	37	634	5.8	162	3,086	5.2	1	51	2.0

[1] Includes centralized telephone and field telephone interviews.
[2] Includes self-administered nonmobile and mobile interviews.
[3] Includes field in-person (CAPI) and field self-administered cases (i.e., sample members who completed the self-administered interview on a field interviewer's [FI] laptop).
NOTE: Admin = The number of sample members to whom an item was administered. Table includes only items that were administered to at least 50 respondents and had declined response rates of 5 percent or greater.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Interview items that asked respondents to report financial information were found to result in relatively high rates of declined response. The items with the overall highest declined-response rates were *Spouse's income—continuous form* (S4INCOMESPS) and *Respondent's income—continuous form* (S4INCOME). Approximately 13 percent of respondents who were administered S4INCOMESP did not respond and 13 percent of those who were administered S4INCOME did not respond. In anticipation that these questions might be left unanswered by many respondents due to their sensitivity, secondary questions that collected income in categorical form (S4INCOMECAT and S4INCOMESPCAT) were also included in the interview, allowing respondents to provide less exact income information. (Declined response rates for S4INCOMECAT and S4INCOMESPCAT were 1 percent and 3 percent, respectively.) *Total amount of private loans for college education* (S4PRVLOANAMT), with a 6 percent declined response rate, also had a follow-up question with categorical choices (S4PRVLOANEST) which allowed respondents to similarly provide a less precise amount or to indicate that they did not know the private loan amount. *Expected yearly salary at age 30* (S4OCC30EARN) was found to have a relatively high rate of declined response, at 7 percent. Though differences in declined response rates between telephone and web modes were relatively small for most items, significantly higher rates of declined response were observed for the telephone administration of many items that collected financial information (S4INCOME: $\chi 2$ (1, N = 17,175) = 157.35, $p < .05$; S4PRVLOANAMT: $\chi 2$ (1, N = 1,964) = 12.42, $p < .05$; S4OCC30EARN: $\chi 2$ (1, N = 16,213) = 373.66, $p < .05$).

Other items found to have relatively high rates of declined response were those that asked respondents to report information about dates. These included *Month ended military service* and *Year ended military service* (S4MLTENDM and S4MLTENDY, at 10 percent and 13 percent, respectively) and *Month last worked for pay before February 2016* and *Year last worked for pay before February 2016* (S4WORKENDM and S4WORKENDY, at 5 percent for each).

The remaining items with declined response rates above 5 percent included *Actively looking for work in February 2016* (S4UNEMP16FB, 8 percent), *February2016/last job duties* (S4JOBDUTY2, 8 percent), and *Ever applied to college* (S4EVERAPPLY, 7 percent). S4JOBDUTY2, *February2016/last job duties*, was one of two items asked to determine a respondent's occupation. An occupation coder application included in the survey used reported job title and job duties to produce a list of occupations from which respondents were asked to select one. The job title has a larger influence than does the job duties over which results are returned by the occupation coder application. Therefore, quality of the occupation codes produced was not greatly affected by omitting these job duties. When responding to the question associated with S4EVERAPPLY, *Ever applied to college,* respondents who had earned college

credits in high school or had enrolled in an adult high school completion program at a college (e.g., a program to prepare for a high school equivalency exam) were instructed to exclude the postsecondary institutions at which they were enrolled or earned credits under these circumstances. This customized wording may have caused confusion about whether a specific institution should be counted as one at which they had ever applied and may have led some respondents to decline to respond.

In summary, few items (11) had high overall rates of declined response in the second follow-up. Additionally, for several of these items, relatively high rates of declined response were anticipated at the outset and follow-up questions included in the survey were designed to reduce missing data.

6.7 Unit and Item Nonresponse Bias Analysis

Unit and item nonresponse bias analyses are presented in this section, with unit nonresponse discussed in section 6.7.1 and item nonresponse discussed in section 6.7.2.

6.7.1 Unit Nonresponse Bias Analysis

NCES Statistical Standard 4-4-1 states that "Any survey stage of data collection with a unit or item response rate less than 85 percent must be evaluated for the potential magnitude of nonresponse bias before the data or any analysis using the data may be released. Estimates of survey characteristics for nonrespondents and respondents are required to assess the potential nonresponse bias" (Seastrom 2014).

The bias in an estimated mean based on respondents $\bar{y}_R$, is the difference between the expected value of this mean and the target parameter, π, the population mean. Analysts can estimate the target parameter for variables that are observed for both respondents (R) and nonrespondents (NR) as follows: $\hat{\pi} = (1-\eta)\bar{y}_R + \eta\bar{y}_{NR}$, where η is the weighted unit (or item) nonresponse rate. For variables that are from the frame rather than from the sample, analysts can estimate π without sampling error. They can then estimate bias as the difference between the respondent mean and the full sample mean: $\hat{B}(\bar{y}_R) = \bar{y}_R - \hat{\pi}$. Equivalently, bias can be estimated as the difference between the mean for respondents and the mean for nonrespondents, multiplied by the weighted nonresponse rate: $\hat{B}(\bar{y}_R) = \eta(\bar{y}_R - \bar{y}_{NR})$. Relative bias provides a measure of the magnitude of the bias relative to the sample mean and is estimated as: $\widehat{RB}(\bar{y}_R) = \hat{B}(\bar{y}_R)/\hat{\pi}$.

Unit nonresponse bias analyses were conducted for the sets of respondents corresponding to the seven analytic weights: the five second follow-up weights and the two supplemental teacher weights for the 2013 Update. Fifteen categorical variables were used to assess unit nonresponse bias. Several of the 15 variables are derived from sampling frame data and are not available in either restricted-use or public-use files. The 15 items are listed below. Variable names are provided for those variables available in a restricted-use file.

- School type (X1CONTROL)
- Charter school status (A1SCHTYPE)
- 9th-grade enrollment by race
- Total school enrollment
- 9th-grade enrollment
- Number of full-time teachers (A1FTTCHRS)
- Student-to-teacher ratio
- Census region (X1REGION)
- School urbanicity (X1LOCALE)
- School grade range (X1GRADESPAN)
- Religious affiliation of school
- Secondary status of school
- State of school (X1STATE)
- Gender (X2SEX)
- Race (X2RACE)

These 15 variables in total comprise 67 categories. The explicit categorization and category labels for each of the 15 items are provided in appendix G. For each category, estimates of bias were calculated and statistical significance tests conducted for each set of respondents corresponding to each of the seven analytic weights.

The results of the nonresponse bias analyses to assess the potential reduction in bias attributable to base weight adjustments for nonresponse are described in the following sections, beginning with a description of the statistical tests for unit nonresponse bias (section 6.7.1.1).

6.7.1.1 Test of nonresponse bias

The VARGEN procedure in SUDAAN was used to estimate bias and conduct *t* tests to determine whether bias was significantly different from zero at a .05 level of significance. Bias estimates were computed for each set of respondents associated with each of the seven analysis weights. For each set of respondents, biases were estimated before weight adjustments were applied to the sampling base weight and then estimated after nonresponse weight adjustments were applied to the sampling base weight. Table 33 contains a summary of the analysis for the five second follow-up and two supplemental 2013 Update analytic weights; see appendix G for the detailed analysis tables. The results of these nonresponse bias analyses suggest that there is not a substantial bias due to nonresponse after adjusting for that nonresponse.

Table 33. Summary statistics for unit nonresponse bias analyses before and after weight adjustments for nonresponse, by HSLS:09 second follow-up and supplemental 2013 Update analytic weights

	Significant bias tests at .05 level[1]		Significant median absolute relative bias tests at .05 level[1,2]		
Analytic weight	Percent before weight adjustment	Percent after weight adjustment	Percent before weight adjustment	Percent after weight adjustment	Percent relative change[3]
[W4STUDENT] Second follow-up	23.9	0	1.8	0	-100.0
[W4W1STU] Base year to second follow-up	25.4	0	2.2	0	-100.0
[W4W1W2W3STU] Base year to first follow-up, to 2013 Update to second follow-up	37.3	0	2.8	0	-100.0
[W4W1STUP1] Base year to second follow-up with base-year parent	38.8	1.5	2.9	0	-100.0
[W4W1STUP1P2] Base year to second follow-up with base-year and first follow-up parent	38.8	1.5	4.9	0	-100.0
[W3W1STUA] Base year to 2013 Update with base-year math teacher	52.2	16.4	7.4	3.3	-55.4
[W3W1STUB] Base year to 2013 Update with base-year science teacher	41.8	7.5	7.2	2.6	-63.9

[1] "Before" and "after" are in reference to the nonresponse weight adjustment. A total of 67 statistical tests were performed; the number 67 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 estimate computed using respondents and nonrespondents. The absolute relative bias is the absolute value of the (percent) relative bias.
[3] The percent relative change is the percentage decrease in median absolute relative bias after weight adjustment. The formula for this was 100 * (median bias value after adjustment – median bias value before adjustment) / median bias value before adjustment.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

6.7.1.2 Second follow-up student-level (W4STUDENT) unit nonresponse bias analysis

In keeping with the NCES statistical standards, nonresponse bias analyses were performed for second follow-up responses using the student analytic weight W4STUDENT because, as shown in table 30, the weighted student response rate for the second follow-up was 67.9 percent. Students who completed a substantial portion of the survey were classified as a respondent; see section 3.5 for further details. Note that participation rates in chapter 4 are based on unweighted cases fielded, and response rates in the current chapter are based on the full sample and are weighted.

Approximately 23.9 percent of the 67 statistical tests conducted for the student-level unit response data identified bias statistically significant at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment, no tests were statistically significant at the .05 level of significance, and the median absolute relative bias was reduced by 100.0 percent. Results of the 67 statistical tests are presented in table G-1 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-2 in appendix G.

6.7.1.3 Base-year to second follow-up student-level (W4W1STU) unit nonresponse bias analysis

As shown in table 30, the weighted unit response rate for the second follow-up was 67.9 percent. However, the weighted unit response rate for students with responses in the second follow-up *and* the base year was 62.5 percent. Approximately 25.4 percent of the 67 statistical tests for this group of respondents identified statistically significant bias at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment, no tests were statistically significant at the .05 level of significance, and the median absolute relative bias was reduced by 100.0 percent. The detailed analyses are shown in table G-3 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-4 in appendix G.

6.7.1.4 Base year to first follow-up to 2013 Update to second follow-up student-level (W4W1W2W3STU) unit nonresponse bias analysis

The weighted unit response rate for student questionnaire responses in the second follow-up, base year, first follow-up, and 2013 Update was 52.0 percent (see table 30). Approximately 37.3 percent of the 67 statistical tests for this group of respondents identified statistically significant bias at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment, no tests were statistically significant at the .05 level of significance, and the median absolute

relative bias was reduced by 100.0 percent. The detailed analyses are shown in table G-5 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-6 in appendix G.

6.7.1.5 Base year to second follow-up with base-year parent student-level (W4W1STUP1) unit nonresponse bias analysis

The weighted unit response rate for student questionnaire responses in the second follow-up and base year, *and* with a responding parent in the base year was 50.1 percent (see table 30). Approximately 38.8 percent of the 67 statistical tests for this group of respondents identified statistically significant bias at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment, bias was found to be statistically different from 0 at the .05 level of significance for 1 of the 67 tests and the median absolute relative bias was reduced by 100.0 percent. The detailed analyses are shown in table G-7 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-8 in appendix G.

6.7.1.6 Base year to second follow-up with base-year and first follow-up parent student-level (W4W1STUP1P2) unit nonresponse bias analysis

The weighted unit response rate for students with responses in the second follow-up and base year, *and* with a responding parent in the base year *and* first-follow-up was 44.6 percent (see table 30). Approximately 38.8 percent of the 67 statistical tests for this group of respondents identified statistically significant bias at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment, bias was found to be statistically different from 0 at the .05 level of significance for 1 of the 67 tests and the median absolute relative bias was reduced by 100.0 percent. The detailed analyses are shown in table G-9 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-10 in appendix G.

6.7.1.7 Base year to 2013 Update with base-year math teacher student-level (W3W1MATHTCH) unit nonresponse bias analysis

The weighted unit response rate for students with responses in the base year and 2013 Update, *and* with a responding math teacher in the base year was 51.4 percent (see table 30). Approximately 52.2 percent of the 67 statistical tests for this group of respondents identified statistically significant bias at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment for

nonresponse and calibration,[74] bias was found to be statistically different from 0 at the .05 level of significance for 11 of the 67 tests and the median absolute relative bias was reduced by 55.4 percent. The detailed analyses are shown in table G-11 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-12 in appendix G.

6.7.1.8 Base year to 2013 Update with base-year science teacher student-level (W3W1SCITCH) unit nonresponse bias analysis

The weighted unit response rate for students with responses in the base year and 2013 Update, *and* with a responding science teacher in the base year was 50.7 percent (see table 30). Approximately 41.8 percent of the 67 statistical tests for this group of respondents identified statistically significant bias at the .05 significance level (see table 33) prior to adjusting the weights for nonresponse. After adjustment for nonresponse and calibration, bias was found to be statistically different from 0 at the .05 level of significance for 5 of the 67 tests and the median absolute relative bias was reduced by 63.9 percent. The detailed analyses are shown in table G-13 in appendix G. Additional comparisons between estimates produced after nonresponse adjustment and estimates produced after poststratification are provided in table G-14 in appendix G.

6.7.2 Item Nonresponse Bias Analysis

NCES Statistical Standard 4-4-3A states: "For an item with a low total response rate, respondents and nonrespondents can be compared on sampling frame and/or questionnaire variables for which data on respondents and nonrespondents are available. Base weights must be used in such analysis. Comparison items should have very high response rates. A full range of available items should be used for these comparisons. This approach may be limited to the extent that items available for respondents and nonrespondents may not be related to the low response rate item being analyzed" (Seastrom 2014).

Moreover, NCES Statistical Standard 1-3-5 states: "Item response rates (RRI) are calculated as the ratio of the number of respondents for whom an in-scope response was obtained (I^x for item x) to the number of respondents who are asked to answer that item. The number asked to answer an item is the number of unit level respondents (I) minus the number of respondents with a valid skip for item x (V^x). When an abbreviated questionnaire is used to convert refusals, the eliminated

[74] In the construction of W3W1MATHTCH, the adjustments for nonresponse and calibration were performed in a single weight adjustment.

questions are treated as item nonresponse. In the case of constructed variables, the numerator includes cases that have available data for the full set of items required to construct the variable, and the denominator includes all respondents eligible to respond to all items in the constructed variable" (Seastrom 2014). The item response rate is calculated as $RRI^x = I^x / (I - V^x)$.

All study items with a weighted response rate (weighted using the second follow-up final analytic weight, W4STUDENT) of less than 85 percent were classified as having high item nonresponse and were included in the item nonresponse bias analyses. These variables and their response rates are described below in section 6.7.2.1.

The procedures for estimating and testing bias are the same as those used for unit nonresponse bias and are described in section 6.7.1. For each study item with less than an 85 percent response rate, as described above, bias estimates are computed by comparing item respondents to all other sample members who were eligible, or assumed eligible, for the item but did not respond to the item. NCES standards require that student questionnaire nonrespondents, whose item eligibility is unknown, must be assumed eligible for the item and must be treated as item nonrespondents. Consequently, bias estimates are computed using the student base weights since these weights are available for student questionnaire nonrespondents. The item nonresponse bias analysis was conducted using a subset of the frame variables used for the unit nonresponse bias analysis. The following school and student characteristics were available for both respondents and nonrespondents from the sampling frame and were used to assess item nonresponse bias:

- School type (X2CONTROL)
- Region of the United States (X2REGION)
- Locale (X2LOCALE)
- Sex (X2SEX)
- Race/ethnicity (X2RACE)

The results of the item nonresponse bias analysis are summarized below in section 6.7.2.2. Detail tables with item-level results appear in appendix G.

6.7.2.1 Variables with high item nonresponse

All second follow-up restricted-use student-level variables were reviewed to identify variables with a response rate below 85 percent. A total of 106 items had a response rate below 85 percent and were included in the nonresponse bias analysis. These variables and their response rates are given in table 34. The lowest weighted item

response rate, 0.6 percent, was found for the variable *S4D12A - Month first adopted child was adopted* (S4ADOPTM).

Item response rates are calculated using both students for whom eligibility is known and students for whom eligibility is not known. Items that have a high completion rate among students with known eligibility may have a relatively small weighted response rate because students with unknown eligibility are assumed to be eligible and treated as nonrespondents. Unknown item eligibility arises for two primary reasons: eligibility for items not administered in the abbreviated questionnaire is unknown for abbreviated questionnaire respondents and eligibility for all items after a gate question is unknown for students who did not answer the gate question.[75]

[75] For sample members who completed the abbreviated questionnaire, eligibility may have been logically inferred for some items not contained in the abbreviated questionnaire when there was a clear correspondence between items. For example, sample members who reported a marital status of "Single and never married" in the abbreviated questionnaire (S4MARITALSTAT) would not have been eligible for the item that collected spouse income (S4INCOMESPS), though it was excluded from the abbreviated data collection instrument.

Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT 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]
S4ALG1WHEN	S4A08 - When took Algebra I	1.6	98.2	0.3	85.6	84.9
S4JOBSAT2	S4C44 - Job satisfaction: February 2016/last job	36.2	57.3	6.5	84.8	84.7
S4APPRENTICE2	S4C45 - Job is apprenticeship: February 2016/last job	36.1	57.3	6.5	84.7	84.6
S4HELPCRSENGL	S4B39E - Requested help for college course: English	40.9	52.4	6.7	85.9	84.6
S4HELPCRSOTH	S4B39F - Requested help for college course: other subject not listed	40.9	52.4	6.7	85.9	84.6
S4WANTXHRS2	S4C43 - Wanted to work more hours: February 2016/last job	36.0	57.3	6.7	84.4	84.3
S4BENHLTH2	S4C38A - Benefits offered in Feb 2016/last job: Health insurance	34.5	59.1	6.4	84.2	84.2
S4JOBDUTY1	S4C22B - Job duties of first job after high school	75.6	11.0	13.4	84.9	84.1
S4BENEDU2	S4C38D - Benefits offered in Feb 2016/last job: Scholarship/tuition reimbursement	34.4	59.1	6.5	84.1	84.1
S4PRVLOANEST	S4B46 - Estimate of total amount of private loans for college education	10.7	87.3	2.1	83.8	84.0
S4BENLIFE2	S4C38B - Benefits offered in Feb 2016/last job: Life insurance	34.4	59.1	6.5	84.1	84.0
S4LASTHSYR	S4A04B - Year last attended high school	7.1	91.6	1.3	84.0	84.0
S4BENVACTN2	S4C38E - Benefits offered in Feb 2016/last job: Paid vacation/sick/personal days	34.4	59.1	6.5	84.0	84.0
S4BENRET2	S4C38C - Benefits offered in Feb 2016/last job: Retirement/financial benefits	34.4	59.1	6.5	84.1	83.9
S4USBORN	S4D33 - Born in the United States	16.4	80.6	3.0	84.6	83.9
S4LASTHSMO	S4A04A - Month last attended high school	7.1	91.6	1.4	83.7	83.8
S4HSPGM16FB	S4A13 - Attending an adult high school completion program in February 2016	1.4	98.3	0.3	81.5	83.7
S4WRK1213	S4C04A - Worked for pay while attending college: 2012–2013 academic year	4.3	94.9	0.8	84.9	83.0

See notes at end of table.

Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT weight—Continued

		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]
S4HELPCRSMTH	S4B39A - Requested help for college course: math	36.3	56.9	6.8	84.3	82.9
S4HSEQUEXPECT	S4A16 - Expects to complete GED or high school equivalency by the end of 2016	4.2	95.0	0.9	82.9	82.8
X4CHOICEAPPID	First choice among colleges applied to	62.9	25.2	11.9	84.1	82.0
X4CHOICEACCID	First choice among colleges accepted to	81.8	1.3	16.9	82.9	82.0
S4OCC30EARN	S4C55 - Expected yearly salary at age 30	57.0	30.6	12.4	82.1	81.9
S4RENTAMT	S4D18 - Amount of housing payment or contribution	34.4	58.8	6.8	83.5	81.7
S4HELPCRSSCI	S4B39B - Requested help for college course: science	27.4	66.4	6.1	81.7	81.5
S4UNEMP16FB	S4C47 - Actively looking for work in February 2016	7.4	91.0	1.6	81.9	81.3
S4TRANSFERACAD	S4B19A - Changed colleges: academic reasons	7.4	91.0	1.6	81.9	81.3
S4TRANSFERFAM	S4B19B - Changed colleges: personal or family reasons	7.4	91.0	1.6	81.9	81.3
S4TRANSFERFIN	S4B19C - Changed colleges: financial reasons	7.4	91.0	1.6	81.9	81.3
S4TRANSFERNONE	S4B19E - Changed colleges: reason not listed	7.4	91.0	1.6	81.9	81.3
S4TRANSFERWRK	S4B19D - Changed colleges: work, military, career reasons	11.7	85.0	3.3	77.9	80.7
S4SPOUSEED	S4D08 - Spouse's/partner's education level	24.8	69.0	6.2	80.0	80.7
S4EVERAPPLY	S4B01 - Ever applied to college	11.7	85.0	3.3	77.8	80.6
S4SPSCLG	S4D06 - Spouse/partner was attending college in February 2016	29.6	63.5	6.9	81.1	80.6
S4CSIMF	S4B35C - Instructors treat male and female students differently: comp sci dept	72.3	12.6	15.1	82.7	80.2
S4PRVLOANAMT	S4B45 - Total amount of private loans for college education	10.1	87.3	2.6	79.5	79.6
S4CSIRC	S4B36C - Instructors treat students of different races differently: computer sci	29.0	63.5	7.5	79.4	78.8

See notes at end of table.

Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT weight—Continued

		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]
S4INCOME	S4D19 - Respondent's income - continuous form	77.6		22.4	77.6	77.3
S4CONTRIBUTE	S4D17 - Contributes to parents' household expenses	38.2	50.4	11.4	77.0	77.2
S4DEPCHILD	S4D23 - Provides more than half of financial support for own child(ren)	8.3	88.6	3.1	73.0	76.9
S4LIVEKIDAMT	S4D14 - Amount of time lives with child(ren)	69.4	12.6	18.0	79.4	76.3
S4BREAKACAD	S4B16A - Took break between high school and college: academic reasons	8.3	88.6	3.1	72.4	76.1
S4BREAKFAM	S4B16B - Took break between high school and college: personal or family reasons	6.4	91.5	2.1	75.3	75.3
S4BREAKFIN	S4B16C - Took break between high school and college: financial reasons	6.4	91.5	2.1	75.3	75.3
S4BREAKNONE	S4B16E - Took break between high school and college: reason not listed	6.4	91.5	2.1	75.3	75.3
S4BREAKWRK	S4B16D - Took break between high school and college: work, military, career	6.4	91.5	2.1	75.3	75.3
X4PARDATE	Date first became parent	6.4	91.5	2.1	75.3	75.3
S4CHILDBORNM	S4D11A - Month first biological child was born	8.2	88.6	3.2	71.6	75.1
S4CHILDBORNY	S4D11B - Year first biological child was born	7.9	88.9	3.2	71.0	74.8
X4IMMIGEN	Immigrant generation	7.8	88.9	3.2	70.8	74.4
S4PARCHILDCR	S4D28A - Parents paid expenses for children or provided childcare	80.2	.	19.8	80.2	74.3
X4ATNDAPPINST	Institution ended up attending as result of first applications	7.8	88.9	3.3	70.5	74.3
S4WRKHRS1213	S4C05A - Hours per week worked while attending college: 2012–2013 academic year	2.1	97.1	0.8	72.9	72.0
S4HELPCRSCSI	S4B39C - Requested help for college course: computer science/technology	18.7	74.6	6.7	73.6	71.5
S4INFORMEDCLG	S4D44 - Ever informed college or trade school of disability or special need	19.6	72.4	8.0	71.0	71.3

See notes at end of table.

Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT weight—Continued

		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]
S4HSEQEXAMPASS	S4A15 - Passed all parts of GED/high school equivalency test the first time	3.4	95.2	1.4	71.5	71.0
S4DEPCHILDNUM	S4D24 - Number of children receive more than half support from respondent	6.0	90.8	3.1	66.0	70.9
S4MARRIAGEY	S4D05B - Year of first marriage	4.2	93.4	2.4	64.0	68.2
S4MARRIAGEM	S4D05A - Month of first marriage	4.2	93.4	2.4	64.1	67.8
S4WRKPGMPAID	S4C02 - Last work experience program paid or unpaid	22.6	67.7	9.7	70.0	66.2
S4WHENAPPLY	S4B02 - When applied to college	11.0	82.8	6.2	63.8	64.6
X4TXSATMATH	College entrance exam math score in terms of SAT	72.7	.	27.3	72.7	64.5
S4ATNDCLGAPP	S4B04 - Attended one of first colleges applied to	11.0	82.8	6.2	63.8	64.5
X4TXACTCOMP	College entrance exam composite score in terms of ACT	72.7	.	27.3	72.7	64.5
X4TXSATCOMP	College entrance exam composite score in terms of SAT	72.7	.	27.3	72.7	64.5
X4TXSATREAD	College entrance exam critical reading score in terms of SAT	72.7	.	27.3	72.7	64.5
S4INCOMESPCAT	S4D22 - Spouse's income - categorical form	3.5	94.1	2.4	59.9	64.3
S4CLGAPPNUM	S4B03 - Number of colleges applied to when first applied	10.9	82.8	6.3	63.6	64.2
S4MHDISBL	S4D39 - Difficulty concentrating/remembering/ deciding due to mental health	19.4	69.2	11.4	63.0	62.8
S4CHOICEAPP	S4B08A - First choice among colleges applied to	10.5	82.8	6.7	61.0	61.4
S4CHOICEAPPID	S4B08B - First choice among colleges applied to - IPEDS ID	10.1	82.8	7.1	59.0	59.5
S4CHOICEACC	S4B11A - First choice among colleges accepted to	9.9	83.2	6.9	58.8	59.0
S4CLGID	S4B05 - IPEDS ID: college attended when first applied	8.7	85.1	6.2	58.2	58.3

See notes at end of table.

Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT weight—Continued

		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]
S4INCOMESPS	S4D21 - Spouse's income - continuous form	3.2	94.1	2.7	54.4	58.0
S4CHOICEACCID	S4B11B - First choice among colleges accepted to - IPEDS ID	9.7	83.2	7.1	57.5	57.7
S4ENGMF	S4B35D - Instructors treat male and female students differently: engineering	9.6	83.5	6.9	58.1	56.4
S4CITIZEN	S4D34 - Citizenship in February 2016	7.0	89.3	3.6	66.0	55.4
S4ENGRC	S4B36D - Instructors treat students of different races differently: engineering	9.4	83.5	7.1	56.8	54.8
S4CLGAPPID1	S4B06 - IPEDS ID: Other college applied to when first applied - 1	6.5	87.2	6.3	50.9	51.3
S4HELPCRSENG	S4B39D - Requested help for college course: engineering	6.6	86.7	6.7	49.5	46.2
S4APPSTATUS1	S4B09 - Outcome of first (other) application	5.7	87.2	7.0	45.0	45.4
S4SPSDEGPGM	S4D07 - Type of degree/certificate spouse/partner working on in Feb 2016	2.3	94.3	3.3	41.0	40.8
S4CLGAPPID2	S4B07 - IPEDS ID: Other college applied to when first applied - 2	3.4	90.4	6.3	34.9	33.7
S4DEPOTHNUM	S4D26 - Number of other dependents	5.0	83.2	11.8	29.5	33.0
S4ACCOMODATION	S4D45 - Received accommodations for disability from any college or trade school	3.6	88.3	8.0	31.2	29.4
S4APPSTATUS2	S4B10 - Outcome of second (other) application	2.6	90.4	7.0	27.1	26.1
S4ACTIVEDUTY	S4C15 - Served on active duty	2.9	87.2	9.9	22.6	23.0
S4MLTGRADE	S4C13 - Highest military pay grade	2.9	87.2	9.9	22.5	23.0
S4MLTSTARTM	S4C09A - Month started military service	2.9	87.2	9.9	22.5	22.9
S4MLTSTARTY	S4C09B - Year started military service	2.9	87.2	9.9	22.4	22.8
S4AIRFORCE	S4C14B - Branch(es) of the military served: Air Force	2.9	87.2	9.9	22.4	22.7

See notes at end of table.

Table 34. Student-level questionnaire items with a weighted item response rate below 85 percent using W4STUDENT weight—Continued

		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]
S4ARMY	S4C14A - Branch(es) of the military served: Army	2.9	87.2	9.9	22.4	22.7
S4COASTGRD	S4C14E - Branch(es) of the military served: Coast Guard	2.9	87.2	9.9	22.4	22.7
S4MARINES	S4C14C - Branch(es) of the military served: Marines	2.9	87.2	9.9	22.4	22.7
S4NAVY	S4C14D - Branch(es) of the military served: Navy	2.9	87.2	9.9	22.4	22.7
S4MLTCOMP	S4C12 - Military component (active duty, Reserves, National Guard)	2.8	87.2	10.0	22.1	22.7
S4STEPPARM	S4D13A - Month first became stepparent	0.8	96.1	3.1	19.7	22.3
S4STEPPARY	S4D13B - Year first became stepparent	0.8	96.1	3.1	19.7	22.3
S4COMBATZN	S4C16 - Served in a combat zone	2.1	88.0	9.9	17.4	17.9
S4OFFERSFIELD	S4B12C - Importance of program offered when choosing first college attended	1.0	92.3	6.7	13.3	12.5
S4COSTATTEND	S4B12B - Importance of cost of attendance when choosing first college attended	1.0	92.3	6.7	13.3	12.4
S4REPUTATION	S4B12A - Importance of academic quality when choosing first college attended	1.0	92.3	6.7	13.3	12.4
S4MLTENDM	S4C11A - Month ended military service	0.4	89.7	9.9	4.0	3.7
S4MLTENDY	S4C11B - Year ended military service	0.4	89.7	9.9	3.9	3.6
S4ADOPTY	S4D12B - Year first adopted child was adopted	0.0	96.8	3.1	1.5	1.3
S4ADOPTM	S4D12A - Month first adopted child was adopted	0.0	96.8	3.1	1.3	0.6

[1] The reserve codes "−7" and "−6" identify the legitimately skipped/not applicable questionnaire items (corresponding to the "Not applicable" column) and "−1", "−3", "−4", "−8", and "−9" identify the questions that should have been answered but were not (corresponding to the "Item missing" column).
[2] Weighted response rates were calculated with the second follow-up student analysis weight (W4STUDENT).
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

6.7.2.2 Item nonresponse bias analysis results

Nonresponse bias results for each item listed in table 34 are included in appendix G. For each item, bias was estimated and tested for each level of the five frame variables used, for a total of 16 estimates per item and 1,696 estimates in total. These estimates were each tested for statistical significance at the .05 significance level. Data in tables 35 and 36 provide summary statistics for bias ratios and relative bias corresponding to each bias estimate. Bias ratios are calculated as the ratio of the estimated bias to the estimated standard error of the bias and provide a measure of the impact of bias on confidence intervals. Relative bias is the ratio of the bias estimate to the full-sample mean (column labeled 'Total' in appendix G tables). This provides a measure of the magnitude of bias that can be compared across the frame variables used in the analysis.

Table 35 summarizes the bias ratios across all bias estimates. Bias ratios larger than 2.0 suggest the effect of item nonresponse may not be negligible. Of the 1,696 bias tests conducted across the 106 student questionnaire items, 43 percent had a bias ratio greater than 2.0.

Table 36 summarizes the significance tests and relative biases for all bias estimates. Overall, 43.8 percent of the bias estimates had statistically significant differences from zero. The average relative bias is -2.1 and the median relative bias is -0.6. The average absolute relative bias which ignores the positive and negative signs on the individual calculations is 17 and the median absolute relative bias is 9.8. The relative bias estimates varied a great deal by frame variable characteristic. For example, the median and average absolute relative biases for private schools are quite high (41.3 and 40.7, respectively) and the estimates for public schools are quite low (2.6 and 2.6, respectively).

Table 35. Frequency distribution of the estimated bias ratios

Study instrument	Range of bias ratio[1]	Frequency[2]	Percent[3]
Student questionnaire[4]	**Total**	1,696	100.0
	0 ≤ bias ratio < 2.0	967	57.0
	2.0 ≤ bias ratio < 5.0	468	27.6
	5.0 ≤ bias ratio	261	15.4

[1] The bias ratio is calculated as the estimated item nonresponse bias divided by the estimated standard error of the bias.
[2] The number of bias ratio calculations falling in the specified range of values.
[3] Percent of bias ratio calculations falling in the specified range of values.
[4] The set of respondents used for bias estimation correspond to the set of students who responded to the second follow-up questionnaire. Such students have a nonzero value for the second follow-up analysis weight W4STUDENT.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Table 36. Summary statistics for student-level item nonresponse bias analyses

Sampling characteristics	Number of *t* tests	Percent of significant *t* tests[1]	Relative bias[2]		Absolute relative bias[3]	
			Average	Median	Average	Median
Total	**1,696**	**43.8**	**-2.1**	**-0.6**	**17.0**	**9.8**
School type						
Public	106	79.2	0.3	0.6	2.6	2.6
Private	106	79.2	-11.0	-10.5	40.7	41.3
Region						
Northeast	106	25.5	-5.3	-3.1	16.8	10.1
Midwest	106	23.6	5.8	5.1	11.2	6.8
South	106	22.6	1.8	0.7	7.7	4.0
West	106	10.4	-4.9	-1.7	9.9	6.3
Locale						
City	106	21.7	6.7	4.6	12.5	6.7
Suburban	106	34.9	-2.4	0.5	10.1	6.7
Town	106	34.9	-8.6	-12.3	14.7	12.5
Rural	106	17.9	-0.8	-0.5	10.6	8.7
Race/ethnicity[4]						
Hispanic	106	41.5	-4.7	-5.3	18.6	15.2
Asian, non-Hispanic	106	47.2	-9.1	-10.5	34.7	24.8
Black, non-Hispanic	106	40.6	-5.4	-7.3	24.8	21.9
Other	106	45.3	2.5	2.3	9.3	6.3
Student sex						
Male	106	87.7	-11.6	-10.0	21.5	18.2
Female	106	87.7	12.4	12.6	25.8	20.0

[1] Percent of *t* tests with $p < .05$.
[2] Relative bias is calculated as 100 times the estimated bias divided by the full-sample mean.
[3] Absolute relative bias is the absolute value of the relative bias.
[4] "Other" includes all other race categories and all undesignated race/ethnicity categories.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up.

Analysts should exercise caution when analyzing items where the results of the item nonresponse bias analysis suggest the presence of nontrivial levels of bias.

6.8 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.

To alleviate the problem of missing data from a respondent record, statistical imputation methods were employed for the second follow-up similar to those used for the HSLS:09 base year, first follow-up, and 2013 Update. Advantages of using imputed values include the ability to use all study respondent records in an 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 for study participants who responded to the HSLS:09 second follow-up. Values were assigned in place of missing responses through single-value imputation or, in the case of composite variables, through derivation from imputed source variables for ten variables; five student questionnaire variables and five composite variables derived from one or more other variables (section 6.8.1). Indicator variables (flags) are included on the analysis file to allow users to easily identify the imputed values. The quality control and evaluative procedures related to imputation are summarized in section 6.8.2.

6.8.1 Imputed Survey Questionnaire Items

Ten key analysis variables were identified for single-value imputation (table 37) from the HSLS:09 second follow-up data. Additional variables were considered for this list but were excluded because of either high item-level response rates or they were deemed to be of lesser analytic importance.

Table 37. Student questionnaire and composite variables included in single value imputation by number and weighted percent of values missing

Student questionnaire and composite variables	Number of values imputed	Weighted percent imputed
Ever attended college by the end of February 2016 (S4EVRATNDCLG)	12	0.03
High school credential status and type, February 2016 (X4HSCOMPSTAT)	33	0.22
Attending college in February 2016 (X4ATNDCLG16FB)	91	0.55
High school credential date (X4HSCOMPDATE)	119	0.88
Had any jobs for pay since high school (S4ANYJOB)	125	0.61
Has child(ren) (S4CHILDREN)	290	1.56
Working for pay in any job in February 2016 (S4WORKING16FB)	310	1.80
Income in 2015 (X4INCOMECAT)	678	4.23
Average hours worked in February 2016 job (X4EMPHRSFB16)	732	4.23
Actively looking for work in February 2016 (S4UNEMP16FB)	1,064	6.11

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

6.8.1.1 Imputation methodology

Stochastic methods were used to impute the missing values for all student questionnaire and composite variables included in table 37. Specifically, a weighted sequential hot-deck (WSHD) statistical imputation procedure (Cox 1980; Iannacchione 1982) using the final student analysis weight associated with the HSLS:09 second follow-up component (W4STUDENT) was applied to the missing values for the variables in table 37 in the order in which they are listed. The WSHD procedure replaces missing data with valid data from a donor record (i.e., 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 in R. 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. 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 is statistically tested (Breiman et al. 1984). The result is a set of imputation classes formed by the partition 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 that were identified through skip patterns in the instrument and literature suggesting an association. The list of variables used as inputs to the CART procedure is provided in table J-1 of appendix J.

Cycling through the imputation variables, that is, the variables that will have imputed values, was part of the imputation process. Once the imputation variables are imputed the first time, the cycle returns and replaces the imputed values for the first imputation variable with the missing code. Then the imputation process re-imputes the first imputed variable using all variables, including the variables with imputed values, on the dataset. Next the imputation process moves to the second imputation variable, replaces the imputed values with missing values, and re-imputes the second variable. This process continues through all the imputation variables and is referred to as the second cycle. There were five cycles implemented for these imputation variables. The reasoning behind the use of cycling is that the imputed values will converge to a reasonable variable.

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).

6.8.1.2 Imputation results

Student questionnaire variables in table 37 are listed in the order in which they 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 6.8.2.

6.8.2 Evaluation of the Imputed Values

After each value was imputed, a set of quality-control checks was implemented to ensure the highest quality of the imputed values. The unweighted distributions of the values before and after the imputation procedure were also compared, both within and across the imputation classes, to identify large areas of change (see table J-2 of appendix J). Differences greater than 5 percent at the .05 significance level were flagged and examined to determine whether changes should be made to the imputation sort or class variables. Finally, data visualizations of value distributions before and after imputation were reviewed for potentially introduced bias.

The imputed variables' distributions within each imputation class were examined to identify classes where imputation might be done in a manner that does not emulate

the raw data distribution. The visualization part is done for the variable in its entirety. Each variable is graphed 3 different ways—raw data, only imputed data, and raw plus imputed data—and compared for indications of introduced bias.

Multivariate consistency checks ensured that relationships among the imputation variables as well as between the imputation variables and key variables used for classification 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 observed 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.

6.9 Disclosure Risk Analysis and Protections

Extensive confidentiality and data security procedures were employed for the HSLS:09 second 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 to external data sources. Because of the paramount importance of protecting the confidentiality of NCES data that contain information about specific individuals, data files were subject to various procedures to minimize disclosure risk. The HSLS:09 second follow-up data products and some of the disclosure treatment methods employed to produce them are described in this section. Details have been excluded from this document to maintain the necessary level of confidentiality.

The disclosure treatment methods used to produce 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.

Some variables were classified as high risk and were globally suppressed in the public-use file. Variables that were globally suppressed in the public-use file had their values set to a reserve code indicating "missing" for all records in the public-use file. Local variable suppression, where some of the values of some variables were suppressed for some, but not all, of the records in the public-use file, was also used

as part of the disclosure protection. Local variable suppression was used to prevent identification of students who were excluded from the HSLS:09 base-year or first follow-up student survey because it was not offered in a format that allowed their meaningful participation (students referred to as "questionnaire incapable" in response status variables) in the base year or first follow-up of HSLS:09 and was also used in certain circumstances to remove rare responses from the public-use file that could potentially be useful for reidentification.

Swapping was applied to certain HSLS:09 second follow-up data items. Swapping was implemented using NCES DataSwap software and utilized specific and 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 identification because every record had some undisclosed probability of having been swapped.

Because perturbation (swapping) of the HSLS:09 second 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.[76] Also, if the analysis determined that the components of a composite variable should be swapped, then the composite variable was recomputed after swapping.

[76] Utility measures include the Pearson contingency coefficient and product-moment correlation, Cramér's V, and Hellinger distance (Gomatam et al. 2005).

Chapter 7. Data File Contents

This chapter provides a concise account of the High School Longitudinal Study of 2009 (HSLS:09) base-year to second follow-up longitudinal data file contents. It addresses the construction and contents of the data file including composite variables and analysis weights.

7.1 Base-year to Second Follow-up Data File

Data produced for the HSLS:09 second follow-up data collection 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 composite variables; see section 7.3 from the HSLS:09 second follow-up student survey instrument as well as all variables included in the base-year, first follow-up, and 2013 Update and High School Transcript data files. Additional variables include those associated with survey-based analysis such as analysis strata and final analysis weights.

The restricted-use student-level file has one record for each student (25,206 records). The public-use file includes only the 23,503 students who responded in either the base year or first follow-up, plus an additional 88 students deemed in scope for the 2013 Update. The number of records in each product is carried forward from the 2013 Update data products; for further details see chapter 7 of *HSLS:09 2013 Update and High School Transcript Data File Documentation* (Ingels et al. 2015).

7.1.1 Restricted-use Data

HSLS:09 second follow-up restricted-use data are available on a DVD that includes restricted-use flat files and an electronic codebook (ECB) application. A license is required to access the restricted-use ECB, and details on obtaining one are available at https://nces.ed.gov/statprog/instruct.asp. The publication number for the restricted-use file is NCES 2018-141. The DVD is available to licensed users at no cost from 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 in the data files; search variable and value names for keywords related to particular

research questions; review the question and item response wording; 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 a 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 program code, including variable and value labels, that can be used with the analyst's own statistical software.

Because HSLS:09 restricted-use data are at a more detailed level—neither recoded nor suppressed—and because the restricted-use data includes additional files (described in section 7.2) which may be linked to the restricted-use student file, for certain analyses analysts may wish to use restricted-use data over public-use or mixed-use data.

7.1.2 Public-use Data

The public-use data are produced from the restricted-use data, the primary difference being that public-use data files do not include the student-institution or student-institution-program level files. Public-use data undergo more restrictive disclosure-avoidance treatment than the restricted-use data, including recoding and suppression as needed. The disclosure treatment developed for the HSLS:09 second follow-up consisted of several steps:

- Review of the collected data and identification of items that may increase risk of disclosure;
- Application of disclosure treatment to the high-risk items to decrease the risk of disclosure;
- Production of restricted-use data files that incorporate the disclosure-treated data; and
- Production of public-use data files, constructed from the disclosure-treated restricted-use files, using additional disclosure limitation methods.

For more details on the disclosure treatment methods used to produce the HSLS:09 second follow-up data files, please see section 6.9.

Online Codebook. HSLS:09 second follow-up data are also available in a public-use version, which is accessible by everyone via the web-based Online Codebook at

https://nces.ed.gov/onlinecodebook. The publication number for the public-use files is NCES 2018-142.

Online Codebook users can explore frequency distributions and tag variables to download from the HSLS:09 public-use dataset. After a set of variables has been tagged for download, the Online Codebook will also create a custom syntax file for use with the user's preferred software package (i.e., SAS, SPSS, Stata, R, S-Plus, or SUDAAN). Alternatively, choosing a generic file format (e.g., ASCII or CSV) allows for the data to be read into most statistical programming language to conduct analyses.

7.1.3 Mixed-use Data Products

DataLab. Data are also available to the general public via the NCES DataLab, found at https://nces.ed.gov/datalab. The DataLab suite of tools, including PowerStats and QuickStats, provide users with access to a combination of public-use variables and a limited number of restricted-use variables. DataLab contains web tools that permit analysis of data without disclosing microdata contents to the user and, as necessary, suppresses or flags estimates that fail to meet reporting standards. DataLab uses the restricted-use version of HSLS:09 second follow-up data, which are neither recoded nor suppressed. However, the second follow-up data contained in DataLab are for the subset of students contained on the public-use data file (23,503 records). QuickStats is the recommended analytic platform for users with only basic statistical needs; the tool allows users to generate simple tables and graphs with a subset of the most frequently used variables (e.g., to summarize college attendance by demographic characteristics). For users with more advanced statistical needs, PowerStats is the recommended analytic tool. PowerStats includes a broader set of variables and enables users to generate complex tables or estimate simple linear or logistic regression models (e.g., to estimate the probability of college attendance using respondent characteristics as independent variables).

7.2 Contents of the Second Follow-up Data Products

The HSLS:09 base-year to second follow-up longitudinal data file student-level file contains the following:

- **Composite variables**
  - base-year student-level composites
  - first follow-up student-level composites
  - 2013 Update student-level composites
  - High School Transcript student-level composites

  - second follow-up student-level composites

- **Questionnaire response data**
  - base-year student questionnaire
  - first follow-up student questionnaire
  - base-year parent questionnaire
  - first follow-up parent questionnaire
  - base-year teacher questionnaire
  - 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
  - 2013 Update questionnaire
  - second follow-up questionnaire

- **Weights**
  - base-year student-level weights
  - first follow-up student-level weights
  - 2013 Update student-level weights
  - High School Transcript student-level weights
  - second follow-up student-level weights
  - Taylor series primary sampling unit (PSU) and stratum identifiers
  - balanced repeated replication (BRR) weights

In addition to the student-level file, the restricted-use data file also contains a base-year school data file that may be used to generate nationally representative estimates of 2009–10 schools with 9th grades. Importantly, school-level data are only representative during the base-year collection, and school-level data collected during the first follow-up are not generalizable to the nation's high schools with 11th grades. First follow-up administrator and counselor questionnaires are available only at the student level because these data apply only to student-level analyses.

The base-year school-level file has not changed since the base-year ECB was produced, and contains

- base-year school-level composite variables and weights;
- base-year administrator questionnaire; and
- base-year counselor questionnaire.

The restricted-use data file contains additional transcript data, specifically

- **High School Transcript school file** provides school-level information (e.g., school type, types of diplomas, and grade scale) related to transcripts and accounts for each school referenced in the high school transcript student course file. This file accounts for all schools, regardless of whether a school provided transcripts.

- **High School Transcript student school file** provides student information (e.g., completion type, reason left school, and transcript-reported grade point average [GPA]) for each school a student attended. This file accounts for all schools linked to a student, regardless of whether the school provided a transcript for the student.

- **High School Transcript school course file** provides school course information (e.g., course name, School Courses for the Exchange of Data [SCED] code, and course attributes for base-year schools that provided course catalogs). The school-course records provide a complete listing of courses offered by the school.

- **High School Transcript student course file** provides student course information (e.g., course name, SCED code, credits earned, and grade received) from transcripts received for each student. These course records are used directly to construct student-level high school transcript composite variables.

Additionally, the restricted-use file contains postsecondary information on the student-institution file and student-institution-program file as described below:

- **Student-institution file** provides information about each college that a student attended or applied to.[77] Indicators are provided to distinguish between institutions attended during and after high school and those that were applied to but not attended. Each record corresponds to a student's institution and survey variable collected for that student at that institution (e.g., date the student started at the institution). Note that a respondent will not be represented on this file if that respondent indicated they applied to, registered at, or attended a postsecondary institution, but provided no additional information.

[77] Applications were limited to three institutions the student most seriously considered attending.

- **Student-institution-program file** provides information about the program the student was enrolled in at the institution. Each record corresponds to a program from a student's institution. This file contains variables such as the type of degree program being pursued by the student and whether or not the student has completed the program. Note that a student-institution combination will not be represented on this file if the respondent provided no additional information about degree program(s) pursued at that institution.

7.3 Variable Naming Schema

Variable names have been developed using the following convention: the first character indicates the data source, the second character indicates the study round, and the remainder is a descriptive name that identifies the variable.

The following list crosswalks the first character with its data source:

- X—composite variables
- W—weights
- S—student questionnaire
- P—parent questionnaire
- A—administrator questionnaire
- C—counselor questionnaire
- M—math teacher questionnaire
- N—science teacher questionnaire
- T—transcripts

The following list crosswalks the second character with its study round:

- 1—base year
- 2—first follow-up
- 3—2013 Update and High School Transcript study
- 4—second follow-up

For questionnaire variables, the variable label provides additional information to link users to the instrument survey specifications and flowcharts. The section of the questionnaire is identified first, followed by the sequential item numbering within the section. Some items have multiple components within the sequential numbering scheme, and the section number receives a letter indicator. For example, S4A04A (Month last attended high school) and S4A04B (Year last attended high school) are the names for items in the student questionnaire (S), second follow-up (4), interview

section A (A), question 4 (04), items A and B (month and year). Appendix K provides a detailed listing of all variable names and corresponding variable labels.

7.4 Reserve Codes

When data are missing at the item- or unit-level, negative-value reserve codes are used to indicate why the item is missing. Table 41 provides a listing of the reserve code values employed in the second follow-up.

Table 38. Reserve code values: 2016

Value	Description
-1	**Item-missing, don't know** Used when a respondent indicated *don't know* as a response to a question.
-3	**Carry-through missing** Used when a respondent does not answer a prerequisite survey question and is therefore not administered the item.
-4	**Item-missing, abbreviated interview** Used for questions that were not administered because an abbreviated version of the questionnaire was administered (see appendix D for the abbreviated survey specifications).
-5	**Data suppressed (public-use data file only)** Used for data that have been suppressed on the public-use file for disclosure reasons.
-6	**Unit-missing, component not applicable** Used, for example, for first follow-up parent survey data when parents were not included in first follow-up parent subsample.
-7	**Item-missing, item not applicable** Used for questions that are not administered because the question in not applicable based on information already known from a prior answer or another data source. For example, sample members who did not attend college will have -7 values for questions about college.
-8	**Unit-missing.** Used for all variables across an entire survey when a sample member did not respond to the survey.
-9	**Item-missing, nonresponse** Used for questions that are not answered within a survey when the respondent was eligible for the question.

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

7.5 Composite Variables

A set of composite variables has been created for each round of HSLS:09. Composite variables—also called derived variables—are usually generated with

responses from two or more questionnaire items, from multiple data sources (e.g., survey and high school transcript), across multiple data collections (e.g., 2013 Update and second follow-up), or from variable recoding, typically for disclosure-avoidance reasons. Some are copied from another source (e.g., a variable supplied in sampling or imported from an external database), and some are new versions of prior-round composites (e.g., family socioeconomic status [SES] using updated occupation prestige scores for parents/guardians [see section 7.5.1]). Composite variable descriptions can be found in appendix L. The HSLS:09 second follow-up data products include composite variables from prior rounds as well as those newly created with data from the second follow-up data collection.

Most of the composite variables can be used as classification variables or independent variables in data analysis. Some of the composites have undergone imputation to address missing responses. Note that all imputed versions of variables have been flagged and are available in composite variables that are named with "_IM" as the variable suffix.[78]

7.5.1 Revised SES indices

Prior-round SES indices incorporated 1989 prestige scores as part of their calculation. However, new prestige scores using 2010 U.S. census codes were recently developed. Further information about prestige scores are available in General Social Survey methodological reports 122 and 124, available at http://gss.norc.org/get-documentation/methodological-reports. In the base year and first follow-up, parent occupations were coded using the 2000 SOC. This coding was performed at the 6-digit SOC level, although, depending on specificity a 2-digit or 3-digit code could have been provided. When an occupation was missing, it was imputed at the 2-digit level.

To provide HSLS:09 data users with revised SES indices,[79] the new prestige scores were linked to occupation codes obtained in prior rounds using the following steps:

1. Recode occupations, which were based on 2000 SOC, to 2010 SOC using the crosswalk.
2. Merge prestige scores onto the Census/SOC crosswalk to provide prestige at the SOC level.

[78] Variables with "_IM" suffixes should not be confused with those ending in "_I." A suffix of "_IM" indicates data that are imputed through statistical procedures (explained in section 6.8), and "_I" indicates data that are edited through logical inferences (explained in section 5.2).
[79] Note that sample members with nonmissing values in base year or first follow-up SES variables have new values for their updated SES variables; the converse is likewise true. Therefore, there was no need to recompute item-level nonresponse for the new SES variables.

3. Produce means for SEI10 (socioeconomic index score using 2010 U.S. census occupation codes) at the 3-digit SOC level and at the 2-digit SOC level.
4. Merge the mean prestige 3-digit value to occupations coded at 3 or more digits.
5. Merge the mean prestige 2-digit value for occupations coded (or imputed) at the 2-digit level.
6. Provide the revised prestige scores as input to SES calculation.

Revised prestige scores, along with household income and parent education, were then used as inputs to calculate revised SES scores.

This page intentionally left blank

References

Adelman, C. (2006). *The Tool Box Revisited: Paths to Degree Completion From High School Through College*. U.S. Department of Education. Washington, DC.

Binder, D.A. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review, 51*(3), 279–292.

Bozick, R., and Ingels, S.J. (2008). *Mathematics Coursetaking and Achievement at the End of High School* (NCES 2008-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Breiman, L., Friedmand, J., Olshen, R., and Stone. C. (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth, Inc.

Chaney, B., Burgdorf, K., and Atash, N. (1997). Influencing Achievement Through High School Graduation Requirements. *Educational Evaluation and Policy Analysis, 19*(3): 229–244.

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.

Cool, V.A., and Keith, T.Z. (1991). Testing a Model of School Learning: Direct and Indirect Effects on Academic Achievement. *Contemporary Educational Psychology, 16*(1): 28–44.

Cox, B.G. (1980). The Weighted Sequential Hot Deck Imputation Procedure. *Proceedings of the Section on Survey Research Methods* (pp. 721–726). Alexandria, VA: The American Statistical Association.

Gomatam, S., Karr, A., and Sanil, P. (2005). Data swapping as a decision problem. *Journal of Official Statistics 21*(4): 635–656.

Groves, R. M., and Heeringa, S. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. *Journal of the Royal Statistical Society Series A: Statistics in Society, 169*(Part 3), 439–457.

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

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. (2013). *High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-Up Data File Documentation* (NCES 2014-361). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ingels, S.J., Pratt, D.J., Herget, D., Bryan, M., Fritch, L.B., Ottem, R., Rogers, J.E., and Wilson, D. (2015). *High School Longitudinal Study of 2009 (HSLS:09) 2013 Update and High School Transcript Data File Documentation* (NCES 2015-036). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Meyer, R.H. (1998). *The Production of Mathematics Skills in High School.* Chicago and Madison, WI: Harris School of Public Policy Studies, The University of Chicago, and Wisconsin Center for Education Research.

National Survey of Student Engagement. (2017). *High-Impact Practices.* Retrieved November 14, 2017 from http://nsse.indiana.edu/html/high_impact_practices.cfm

Peytchev, A. (2013). Consequences of Survey Nonresponse. *Annals of the American Academy of Political and Social Science, 645*(1): 88–111.

R Core Team. (2015). *R*: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved December 12, 2015 from https://www.R-project.org/.

Ripley, B. (2015). *tree:* Classification and Regression Trees. R package version 1.0-35. Retrieved April 14, 2015 from https://CRAN.R-project.org/package=tree.

Rock, D.A., and Pollack,J.M.. (1995). *Mathematics Coursetaking and Gains in Mathematics Achievement* (NCES 95-714). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Seastrom, M.M. (2014). *2012 Revision of NCES Statistical Standards.* (NCES 2014-097). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

StandardsWork. (2006). *12th Grade NAEP Revised: What Stakeholders Say*. Prepared for National Assessment Governing Board, National Assessment of Educational Progress. Washington, DC: Author.

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.